DOCUMENT RESUME

ED 385 959 EC 304 063

AUTHOR Canivez, Gary L.
TITLE Concurrent Validity and Diagnostic Efficiency of the Kaufman Brief Intelligence Test in Assessing Severe Discrepancy in Reevaluations of Students with Learning Disabilities.

PUB DATE Mar 94
NOTE 38p.; Paper presented at the Annual Meeting of the National Association of School Psychologists (26th, Seattle, WA, March 4-5, 1994).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Concurrent Validity; *Diagnostic Tests; Efficiency; Elementary Education; Elementary School Students; *Intelligence Tests; *Learning Disabilities; Measures (Individuals); Middle Schools; Student Evaluation; *Test Validity
IDENTIFIERS *Ability Achievement Discrepancy; *Kaufman Brief Intelligence Test; Middle School Students; Wechsler Intelligence Scale for Children III; Woodcock Johnson Revised Tests of Achievement

ABSTRACT

This study examined the concurrent validity and diagnostic efficiency of the Kaufman Brief Intelligence Test (K-BIT) with 75 elementary and middle school students with learning disabilities, who had been referred for triennial multidisciplinary re-evaluations. High and significant correlations were found between the K-BIT and the Wechsler Intelligence Scale for Children-III (WISC-III). However, K-BIT Vocabulary-Matrices discrepancies were not in agreement with WISC-III Performance IQ-Verbal IQ discrepancies. Substantial agreement was found between the presence or absence of achievement-ability discrepancies identified by the K-BIT Composite and the Woodcock-Johnson Revised Tests of Achievement (WJ-R ACH) and those identified by the WISC-III full scale IQ and the WJ-R ACH. Results suggest that diagnostic precision can be maintained and time may be saved by utilizing a brief ability measure (K-BIT) as a substitute for a comprehensive ability measure (WISC-III) in the re-evaluation process of students with learning disability. (Contains 16 tables and 29 references.) (Author/DB)

Concurrent Validity and Diagnostic Efficiency of the Kaufman Brief Intelligence Test in Assessing Severe Discrepancy in Reevaluations of Students with Learning Disabilities

Gary L. Canivez

Eastern Illinois University

Gary L. Canivez is an assistant professor in the Department of Psychology at Eastern Illinois University.

The 75 subjects previously classified with learning disabilities in this study are a subset of subjects from a larger study (*N* = 137; Canivez, 1995) addressing the concurrent validity of the K-BIT which was also presented at the 1994 Annual Convention of the National Association of School Psychologists, Seattle, WA. This manuscript is presented as an ERIC document in order to present *all* tables pertaining to diagnostic efficiency of the K-BIT which is not possible in journals due to space limitations.

I thank W. Ken Newman and John Wilson for assistance in data collection and comments on the previous versions of this manuscript. I also thank Joseph C. Kush and Marley W. Watkins for comments on previous versions of this manuscript.

Correspondence concerning this article should be addressed to Gary L. Canivez, Ph.D., Department of Psychology, Eastern Illinois University, 600 Lincoln Avenue, Charleston, IL 61920. An operational version of the *Microsoft® Excel*™ spreadsheet template may be obtained from Dr. Canivez by sending an *Apple® Macintosh*™ formatted 3.5 inch disk with a self-addressed, stamped envelope and $5.00 to cover duplication expenses.

Abstract

The present study examined the concurrent validity and diagnostic efficiency of the K-BIT. Correlations between K-BIT and WISC-III were quite high and significant (except WISC-III PSI). K-BIT Vocabulary - Matrices discrepancies were not in agreement with WISC-III VIQ - PIQ discrepancies resulting in nonsignificant κ coefficients for various significance levels. Substantial agreement was found between the presence or absence of achievement-ability discrepancies identified by the K-BIT IQ Composite and WJ-R ACH and those identified by the WISC-III FSIQ and WJ-R ACH. These results suggest that time may be saved by utilizing a brief ability measure (K-BIT) as a substitute for a comprehensive ability measure (WISC-III) in the reevaluation process of students with learning disability and retain a high degree of diagnostic precision.

Concurrent Validity and Diagnostic Efficiency of the Kaufman Brief Intelligence Test in Assessing Severe Discrepancy in Reevaluations of Students with Learning Disabilities

School psychologists report spending significant portions of time (1/2 to 2/3) in evaluation of students for possible placement in special education programs (Goh, Teslow, & Fuller, 1981; Hutton, Dubes, & Muir; 1992; Reschly, Genshaft, & Binder, 1987; Smith, 1984). Specific learning disability (SLD) has become the category of special education with the highest proportion of students (Heath & Kush, 1991); thus, much of a school psychologist's time is spent evaluating such students (Reschly et al., 1987). In the evaluation and identification of students with SLD it is necessary, in part, to identify the presence of "a severe discrepancy between achievement and intellectual ability" (United States Department of Education [USDE], 1992, p. 44823). Individually administered comprehensive intellectual ability measures are most frequently used to assess intellectual ability; and the Wechsler scales are by far the most frequently used among school psychologists (Goh et al., 1981; Hutton et al., 1992).

In assessing the presence of severe discrepancy between achievement and intellectual ability, several experts in psychological measurement indicate that the technically appropriate method for determining severe discrepancy between achievement and intellectual ability (when both instruments were not co-normed) is through the use of a regression based mathematical formula (Heath and Kush, 1991; Reynolds, 1984; Wilson & Cone, 1984). This approach accounts for regression to the mean effects as well as measurement error. This approach may also assist in reducing error in identifying students as learning disabled (Telzrow, 1990).

Once identified as disabled, reevaluation is required "every three years, or more frequently if conditions warrant" (USDE, 1992, p.44822). In the reevaluation process, there is no specification (or mandate) to replicate previously used instruments, although this usually occurs. In the reevaluation of students with SLD another comprehensive intellectual assessment is typically provided. Given time constraints, the readministration of a comprehensive intellectual measure may not be time or cost effective practice, particularly if the test yields relatively unchanged ability estimates. The use of an intellectual screening test to recheck the intellectual status of the referred student could save time that might be better spent evaluating the effectiveness of the individual education program (Ross-Reynolds, 1990) or in the provision other types of services (e.g., consultation, counseling, research, and program development).

One of the recommended uses of the Kaufman Brief Intelligence Test (K-BIT) is rechecking the intellectual status of an individual when they had previously been administered a comprehensive intelligence test (Kaufman & Kaufman, 1990). In addition the K-BIT was also developed to measure and compare verbal and nonverbal abilities as is done with the Wechsler scales. Although studies to date indicate a high degree of concurrent and convergent validity when comparing the K-BIT to comprehensive ability measures (Naugle, Chelune, & Tucker, 1993; Kaufman & Kaufman, 1990; Prewett, 1992a, 1992b) little is known about the psychometric relations between the K-BIT and the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) (Wechsler, 1991) as there is to date only one published study available (Canivez, 1995). In addition, there also are no published studies to date investigating the validity of the K-BIT among a sample of students with SLD. The use of the K-BIT in the reevaluation of students with SLD requires careful study to determine if it yields similar information and diagnostic power when compared to a comprehensive measure such as the WISC-III.

The present study investigated the concurrent validity of the K-BIT in comparison with the WISC-III in a sample of students previously identified with SLD. Examination of the diagnostic utility of the K-BIT in identifying the presence or absence of severe achievement-ability discrepancies when compared to the presence or absence of severe ability-achievement discrepancies identified by the WISC-III was of particular importance. If the K-BIT is to be a useful instrument in reducing the reevaluation time of students with SLD by supplanting a comprehensive intellectual ability measure, it must have acceptable levels of diagnostic agreement when compared to results obtained from a comprehensive intellectual ability measure.

Method

Subjects

The 75 subjects in the present study were elementary (K-6) and middle school (6-8) students in a major southwest metropolitan public school system who were referred for triennial multidisciplinary reevaluations. All were previously identified as specific learning disabled according to state special education rules and regulations used for classification of students as learning disabled which were similar to those specified by the United States Department of Education (1992). Learning disability was operationally defined as a severe discrepancy between ability and achievement using a regression approach and 1.5 standard errors of estimate suggested as a minimum criterion for severe discrepancy when the students were initially classified by multidisciplinary evaluation teams. Sixty-five percent (n = 49) were male, 35% (n = 26) were female, and the mean age of the subjects was 11.79 years (*SD* = 2.04, range = 6-15 years). Ethnic characteristics of the subjects were as follows: Caucasian, 35% (*n* = 26); Black, 9% (*n* = 7); Hispanic, 44% (*n* = 33); Native American, 11% (*n* = 8); and Hispanic-Native American, 1% (*n* = 1). All subjects in this study were sufficiently proficient in English for appropriate administration of present instruments, although some were bilingual. Bilingual subjects were evaluated by a bilingual school psychologist. Sixty-nine percent (*n* = 52) were monolingual English speakers while 17% (*n* = 13) had primary language of English and secondary language of Spanish, and 13% (*n* = 10) had primary language of Spanish and secondary language of English.

Measures

Kaufman Brief Intelligence Test. "The Kaufman Brief Intelligence Test (K-BIT) is a brief, individually administered measure of the verbal and nonverbal intelligence of a wide range of children, adolescents, and adults, spanning the ages of 4 to 90 years" (Kaufman & Kaufman, 1990, p. 1). It is comprised of two subtests: Vocabulary (Expressive Vocabulary and Definitions) and Matrices; and takes approximately 15 to 30 minutes to administer. The K-BIT was standardized on a representative sample (*n* = 2,022) closely approximating 1990 United States Census data on variables of gender, geographic region, socioeconomic status, and race/ethnic group.

Split-half internal consistency reliability estimates across the entire age range for the K-BIT IQ Composite, Vocabulary, and Matrices scores were high; ranging from .88 to .98 (M_r = .94), .89 to .98 (M_r = .93), and .74 to .95 (M_r = .88), respectively. Test-retest stability estimates for the IQ Composite, Vocabulary, and Matrices scores with four age samples ranged from .92 to .95 (M_r = .94), .86 to .97 (M_r = .94), and .80 to .92 (M_r = .85), respectively (Kaufman & Kaufman, 1990).

Concurrent validity studies between the K-BIT and the Test of Nonverbal Intelligence (TONI) yielded low correlations but correlations with the Slosson Intelligence Test (SIT) were moderately high across 5 samples (Kaufman & Kaufman, 1991). Correlation coefficients indicated good support for concurrent validity with the Kaufman Assessment Battery for Children (K-ABC), Wechsler Intelligence Scale for Children-Revised (WISC-R), and Wechsler Adult Intelligence Scale-Revised (WAIS-R) for normal samples. K-BIT IQ Composite scores correlated between .58 and .69 (M_r = .63) with the K-ABC Mental Processing Composite across three age ranges. Correlations between the K-BIT IQ Composite and WISC-R Full Scale IQ (r = .80) and WAIS-R Full Scale IQ (r = .75) also strongly supported the construct validity of the K-BIT. In addition to composite score correlations, The K-BIT Vocabulary subtest correlated better with the WISC-R and WAIS-R Verbal IQ's (as would be expected) while the K-BIT Matrices subtest correlated equally well with the WISC-R and WAIS-R Verbal and Performance IQ's (Kaufman & Kaufman, 1990). Prewett (1992a, 1992b) has also found significant correlations between the K-BIT and WISC-R for samples of referred students and incarcerated juvenile delinquents while Naugle, et al., (1993) found significant correlations with the WAIS-R for a sample of patients receiving neuropsychological evaluations.

Wechsler Intelligence Scale for Children-Third Edition. The Wechsler Intelligence Scale for Children-Third Edition (WISC-III) is an individually administered test of intellectual abilities for children aged 6 years through 16 years, 11 months (Wechsler, 1991). As with previous editions, the WISC-III is comprised of several subtests which measure different aspects of intelligence and yields three composite IQ's (viz., Verbal [VIQ], Performance [PIQ], and Full Scale [FSIQ]) which provide estimates of the individual's verbal, nonverbal, and general intellectual abilities. The WISC-III also yields four optional factor-based index scores (viz., Verbal Comprehension [VCI], Perceptual Organization [POI], Freedom from Distractibility [FDI], and Processing Speed [PSI]). The WISC-III was standardized on a representative sample (n = 2,200) closely approximating the 1988 United States Census on gender, parent education (SES), race/ethnicity, and geographic region. Internal consistency reliability estimates for the 3 IQ and 4 Index scores were excellent, ranging from .80 to .97 within the 11 age levels with 55 of 77 (71%) coefficients ≥ .90. Average test-retest stability estimates for the 3 IQ and 4 Index scores were also excellent, ranging from .82 to .94. Concurrent validity studies generally found moderately high correlations with other intellectual ability measures and VIQ tended to correlate higher with verbal ability measures than nonverbal ability measures while PIQ tended to correlate higher with nonverbal ability measures than verbal ability measures (Wechsler, 1991) as expected.

Woodcock-Johnson-Revised Tests of Achievement. The Woodcock-Johnson-Revised Tests of Achievement (WJ-R ACH) is an individually administered test of academic achievement assessing various aspects of reading, mathematics, writing, and general knowledge. Achievement subtests in the Standard Battery (viz., Letter-Word Identification, Passage Comprehension, Calculation, Applied Problems, Dictation, Writing Samples, Science, Social Studies, and Humanities) combine to form four achievement clusters (viz., Broad Reading, Broad Mathematics, Broad Written Language, and Broad Knowledge). Internal consistency reliability coefficients were high with *Mdn* correlations ranging from .87 to .93 across the entire age range for the Standard Battery. Concurrent validity data presented in the *Examiners Manual* (Woodcock & Mather, 1989) indicated that the WJ-R ACH clusters correlated moderately well with other measures of academic achievement assessing similar domains.

Procedure

Subjects were administered the K-BIT, WISC-III, and WJ-R ACH as part of comprehensive triennial multidisciplinary reevaluations. The K-BIT and WISC-III were administered in counterbalanced order, during the same test session, by one of three licensed and nationally certified school psychologists. The WJ-R was administrated, in most instances, by the student's special education teacher, however; for some subjects, the WJ-R was administered by the school psychologist. K-BIT Vocabulary, Matrices, and IQ Composite standard scores were obtained and Vocabulary - Matrices discrepancy scores were evaluated for significant differences at the $\alpha = .05$ and $\alpha = .01$ levels (see Table C.5, Kaufman & Kaufman, 1990, p. 112). Vocabulary - Matrices discrepancy scores were also evaluated for "abnormality" based upon a 5% population prevalence criteria (see Table 3.2, Kaufman & Kaufman, 1990, p. 46).

WISC-III VIQ, PIQ, FSIQ, VCI, POI, FDI, and PSI scores were obtained but of the 75 subjects, 2 were not administered the Symbol Search subtest. Thus, analyses for the Processing Speed Index are based on $n = 73$. VIQ - PIQ discrepancy scores were evaluated for significance for $\alpha = .05$ (see Table B.1, Wechsler, 1991, p. 261) and $\alpha = .01$. Critical values for VIQ-PIQ significance for $\alpha = .01$ are not available in the *WISC-III Manual*, and while Naglieri (1993) provided critical values for significant VIQ - PIQ differences ($\alpha = .01$), these values are inflated for use here due to Bonferroni correction which adjusts for the familywide error rate in multiple discrepancy comparisons. The present study examined only one WISC-III pairwise comparison (viz., VIQ - PIQ), so critical values for significance for $\alpha = .01$ were obtained following the formula:

$$\text{Difference Score} = z(SEM_a^2 + SEM_b^2)^{1/2},$$

where $z = 2.5758$ (value from the normal curve corresponding to $\alpha = .01$), SEM_a = standard error of measurement for VIQ for the appropriate age level, and SEM_b = standard error of measurement for PIQ for the appropriate age level (Anastasi, 1988; Guilford & Fruchter, 1978). The *SEM*s used for each age level were obtained from Table 5.2 in the *WISC-III Manual* (Wechsler, 1991, p. 168). VIQ - PIQ discrepancies were also considered "abnormal" at or below the 5% population prevalence criterion level (see Table B.2, Wechsler, 1991, p. 262).

Raw scores from the WJ-R ACH were converted to standard scores ($M = 100$, $SD = 15$) based upon *age* norms. Most students were administered all WJ-R ACH subtests, however; some were only administered subtests related to their suspected disabilities.

Pearson product-moment correlation coefficients were calculated between the K-BIT Vocabulary, Matrices, and IQ Composite standard scores and the WISC-III VIQ, PIQ, FSIQ, VCI, POI, FDI, and PSI scores. In addition, the K-BIT Vocabulary - Matrices discrepancy score was used as a predictor (continuous independent variable) of the WISC-III VIQ - PIQ discrepancy score (dependent variable) in a linear regression analysis. Diagnostic efficiency statistics were calculated as recommended by Kessel and Zimmerman (1993) to further evaluate the K-BIT Vocabulary - Matrices discrepancy. *Kappa* (κ) coefficients (Cohen, 1960) were calculated to assess the degree of agreement between Vocabulary - Matrices and VIQ - PIQ discrepancies for $\alpha = .05$, $\alpha = .01$ and for the 5% population prevalence criterion. To test whether κ coefficients were significant, z-tests were performed as recommended by

Fleiss (1981, p. 219) and Fleiss, Cohen, & Everitt (1969). Diagnostic efficiency statistics, significance tests, and probability levels were calculated using a modified version of a *Microsoft® Excel™* spreadsheet template (Canivez & Watkins, in press) to eliminate calculation errors. The standard 2 X 2 diagnostic efficiency table was modified to a 3 X 3 table to accommodate the three possibilities of verbal-nonverbal ability discrepancy results (viz., not significant, VIQ/Vocabulary > PIQ/Matrices, or PIQ/Matrices > VIQ/Vocabulary). Pearson product-moment correlation coefficients were also obtained between the K-BIT and WJ-R ACH and between the WISC-III and WJ-R ACH.

The primary investigation in the present study was determining the level of agreement between severe achievement-ability discrepancies identified between subjects' predicted achievement based on their K-BIT IQ Composite scores and actual achievement on WJ-R ACH subtests with severe achievement-ability discrepancies identified between subjects' predicted achievement based on their WISC-III FSIQ scores and actual achievement on WJ-R ACH subtests. Predicted achievement was obtained using the formula:

$$\text{Predicted Achievement} = r_{xy}(IQ - M_{IQ}) + M_{ACH},$$

where IQ = the obtained IQ score, $M_{IQ} = 100$ (average IQ score), and $M_{ACH} = 100$ (average achievement score). Because the actual IQ-Achievement correlations in the present study are likely to underestimate the true relationships in the general population due to restricted range and because the relationships between the WISC-III and WJ-R ACH and K-BIT and WJ-R ACH in the general population were not known; $r_{xy} = .65$ (Heath & Kush, 1991). Severe discrepancy between predicted achievement and actual achievement was defined by the formula:

$$D > 15z(1 - r_{xy})^{1/2}$$

(Reynolds, 1984), where D = Predicted Achievement - Actual Achievement, $z = 1.65$ (z corresponding to $\alpha = .05$ in a one-tail significance test), and $r_{xy} = .65$ (median IQ-Achievement correlation recommended by Heath & Kush, 1991).

Diagnostic efficiency tables comparing the presence or absence of severe achievement-ability discrepancies between the K-BIT predicted achievement and WJ-R ACH with the presence or absence of severe achievement-ability discrepancies between the WISC-III predicted achievement and WJ-R ACH were created using a *Microsoft® Excel™* spreadsheet template (Canivez & Watkins, in press) to eliminate calculation errors. The spreadsheet also calculated all diagnostic efficiency statistics recommended by Kessel and Zimmerman (1993). To test whether κ coefficients were significant, z-tests were performed as recommended by Fleiss (1981, p. 219) and Fleiss et al. (1969).

Results and Discussion

Concurrent Validity

Pearson product-moment correlation coefficients and r^2's for the K-BIT and WISC-III are presented in Table 1. All correlations (except PSI) were significant ($p < .0001$). Correlations ranged from .18 to .82 ($M_r = .62$) The magnitude of these results was somewhat surprising given the restricted range clinical samples normally yield. Judging from the *SD*s and range from the WISC-III and K-BIT, this sample did indeed have a restricted range compared to the standardization samples. Consistent with previous investigations between the K-BIT and WISC-R (Kaufman & Kaufman, 1990; Prewett, 1992a, 1992b) and WAIS-R (Naugle, Chelune, & Tucker; 1993), the K-BIT IQ Composite correlated significantly with the WISC-III FSIQ ($r = .82$) and 67% of the variability of FSIQ was accounted for by the K-BIT IQ Composite. Differences between correlation coefficients were tested using Hotelling's formula for a t-test when coefficients of correlation are correlated (Guilford & Fruchter, 1978, p. 164). As expected, the Vocabulary subtest had a significantly higher correlation with the WISC-III VIQ than with PIQ, $t(72) = 3.32$, $p < .001$, and significantly higher correlation with VCI than with POI, $t(72) = 2.89$, $p < .005$. The Matrices subtest correlated equally well with

Table 1

Pearson product-moment correlation coefficients between the K-BIT and WISC-III (n = 75)

	K-BIT		
	Vocabulary	Matrices	IQ Composite
WISC-III			
VIQ	.72 (.51)	.60 (.36)	.81 (.66)
PIQ	.51 (.26)	.64 (.41)	.71 (.50)
FSIQ	.67 (.45)	.67 (.45)	.82 (.67)
VCI	.73 (.53)	.56 (.31)	.78 (.61)
POI	.55 (.30)	.62 (.38)	.72 (.52)
FDI	.48 (.23)	.56 (.31)	.64 (.41)
PSI[a]	.18 (.03)*	.26 (.07)**	.27 (.07)**

Note. K-BIT = Kaufman Brief Intelligence Test; WISC-III = Wechsler Intelligence Scale for Children-Third Edition; VIQ = Verbal IQ; PIQ = Performance IQ; FSIQ = Full Scale IQ; VCI = Verbal Comprehension Index; POI = Perceptual Organization Index; FDI = Freedom from Distractibility Index; PSI = Processing Speed Index. All correlations significant $p < .0001$ except where noted. r^2's presented in parentheses.
[a]$n = 73$. *ns. **$p < .05$.

FIQ, POI, VIQ and VCI as no significant differences were noted among the correlations. These correlational results were similarly obtained in other studies (Kaufman & Kaufman, 1990; Prewett, 1992a, 1992b). The lowest correlations were with the PSI as expected.

Table 2

Descriptive statistics for K-BIT and WISC-III scores (n = 75)

	M	*SD*	Range
WISC-III			
VIQ	78.05	13.08	55 - 113
PIQ	90.65	13.59	62 - 126
FSIQ	82.68	13.07	59 - 119
VCI	79.17	13.39	52 - 108
POI	92.20	14.82	60 - 131
FDI	79.77	10.67	61 - 112
PSI[a]	92.82	12.61	64 - 122
K-BIT			
Vocabulary	79.49	13.16	44 - 102
Matrices	86.65	14.37	56 - 130
IQ Composite	81.27	12.32	57 - 108

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition, VIQ = Verbal IQ, PIQ = Performance IQ, FSIQ = Full Scale IQ, VCI = Verbal Comprehension Index, POI = Perceptual Organization Index, FDI = Freedom from Distractibility Index, PSI = Processing Speed Index, K-BIT = Kaufman Brief Intelligence Test.

[a]*n* = 73

Descriptive statistics for the K-BIT and WISC-III are presented in Table 2. Subjects' K-BIT IQ Composite and WISC-III Full Scale IQ scores did not differ, $t(74) = 1.62$, *ns*, nor did their K-BIT Vocabulary subtest and WISC-III VIQ scores, $t(74) = 1.28$, *ns*. However, subjects obtained significantly lower K-BIT Matrices subtest scores than WISC-III PIQ scores, $t(74) = 2.91$, $p < .005$; results also found with the WISC-R (Prewett, 1992a, 1992b). Although significant, these mean differences were not large or of practical significance as they are well within the standard errors of measurement for both measures. Naugle, et al., (1993) also reported significant but small differences with subjects scoring consistently *higher* on the K-BIT.

The regression analysis assessing the ability of the K-BIT Vocabulary - Matrices discrepancy score to predict the WISC-III VIQ - PIQ discrepancy score was significant $F(1, 73) = 5.68$, $p < .02$. However, only 7% of the variability in WISC-III VIQ - PIQ discrepancies was accounted for by K-BIT Vocabulary-Matrices discrepancies. Naugle et al., (1993) found the K-BIT Vocabulary - Matrices discrepancy accounted for only 21% of the variability in WAIS-R VIQ - PIQ discrepancies. VIQ - PIQ discrepancies ($M = -12.60$, $SD = 10.33$) were also significantly larger than Vocabulary - Matrices discrepancies ($M = -7.16$, $SD = 16.03$), $t(74) = 2.84$, $p < .006$. Table 3 presents frequency data for students demonstrating various K-BIT Vocabulary - Matrices discrepancies and WISC-III VIQ - PIQ discrepancies for $\alpha = .05$ and .01 and for the 5% population prevalence level. Table 4 presents the diagnostic efficiency statistics for these comparisons.

These data indicated that for $\alpha = .05$ and .01 and for the 5% population prevalence level, κ coefficients were not significant and represented chance

Table 3

Frequencies of subjects showing significant (α = .05 and .01) and "Abnormal" (≤5% population prevalence) K-BIT Vocabulary - Matrices and WISC-III Verbal IQ - Performance IQ discrepancies

α = .05	WISC-III VIQ - PIQ		
K-BIT Vocabulary - Matrices	*ns*	VIQ > PIQ	PIQ > VIQ
ns	15_d	0_c	18_c
Vocabulary > Matrices	9_b	0_a	2
Matrices > Vocabulary	11_b	0	20_a

α = .01	WISC-III VIQ - PIQ		
K-BIT Vocabulary - Matrices	*ns*	VIQ > PIQ	PIQ > VIQ
ns	30_d	0_c	20_c
Vocabulary > Matrices	6_b	0_a	0
Matrices > Vocabulary	11_b	0	8_a

≤5% Population Prevalence	WISC-III VIQ - PIQ		
K-BIT Vocabulary - Matrices	> 5%	VIQ > PIQ	PIQ > VIQ
> 5%	56_d	0_c	10_c
Vocabulary > Matrices	1_b	0_a	0
Matrices > Vocabulary	8_b	0	0_a

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; VIQ = Verbal IQ; PIQ = Performance IQ; K-BIT = Kaufman Brief Intelligence Test. Numbers along the diagonal indicate consistent results and agreement between K-BIT Vocabulary - Matrices discrepancy and WISC-III VIQ - PIQ discrepancy. False negatives fall above the diagonal while false positives fall below the diagonal. Subscripts a, b, c, and d correspond to the appropriate cells in a 2 X 2 diagnostic efficiency statistics table presented in Kessel and Zimmerman (1993).

Table 4

Diagnostic efficiency statistics for agreement between K-BIT Vocabulary - Matrices and WISC-III Verbal IQ - Performance IQ discrepancies

	α = .05	α = .01	5% PP
Sensitivity	.53	.29	.00
Specificity	.43	.64	.86
Positive Predictive Power	.50	.32	.00
Negative Predictive Power	.45	.60	.85
False Positive Rate	.57	.36	.14
False Negative Rate	.47	.71	1.00
Overall Correct Classification	.47	.51	.75
κ	.07	-.01	-.14
SE_{κ}	.09	.10	.11
z	.77	-.12	-1.23
p	*ns*	*ns*	*ns*

Note. PP = Population Prevalence. *ns* = not significant.

levels of agreement between Vocabulary - Matrices and VIQ - PIQ discrepancies. Interestingly, for α = .05, two subjects, showed a significant K-BIT Vocabulary > Matrices discrepancy but demonstrated a significant WISC-III PIQ > VIQ discrepancy (opposite of predicted direction).

Given the small proportion of variability of WISC-III VIQ - PIQ discrepancies accounted for by K-BIT Vocabulary - Matrices discrepancies (7%), low sensitivity estimates, low positive predictive power, and the high false positive and false negative predictions from the K-BIT Vocabulary -

Matrices discrepancy, clinicians should not use the K-BIT Vocabulary - Matrices discrepancy to make predictions or formulate hypotheses regarding possible verbal-nonverbal differences in comprehensive intelligence tests such as the WISC-III (or WAIS-R, Naugle et al., 1993). This may be partly related to the fact that the K-BIT is only comprised of two subtests and does not sample the respective domains as well as a comprehensive intellectual measure. It may also be due to the unreliability seen in difference (discrepancy) scores (Thorndike & Hagen, 1977).

Diagnostic Agreement

Tables 5 and 6 present correlation coefficients between the K-BIT and WJ-R ACH and WISC-III and WJ-R ACH, respectively. Correlations are generally lower (but still significant) than those found between the K-BIT and WISC-III and provide evidence for construct validity as individual intelligence and achievement tests (heterotrait-monomethod) are designed to measure somewhat different domains (Campbell & Fiske, 1959). Verbal ability estimates (Vocabulary and VIQ) correlated as high or higher with WJ-R ACH than general intellectual ability estimates (IQ Composite and FSIQ). It is also interesting to note that nonverbal ability measures (Matrices and PIQ) yielded lower correlations with WJ-R ACH than verbal ability measures (Vocabulary and VIQ) and general intellectual ability estimates (IQ Composite and FSIQ) as expected (Kaufman & Kaufman, 1990; Wechsler, 1991).

Table 5

Pearson product-moment correlation coefficients between the K-BIT and WJ-R ACH

		K-BIT		
	n	Vocabulary	Matrices	IQ Composite
WJ-R ACH				
LWID	67	.48 (.23)****	.17 (.03)	.39 (.15)***
PC	67	.58 (.34)****	.36 (.13)**	.57 (.32)****
C	66	.45 (.20)***	.25 (.06)*	.42 (.18)***
AP	66	.50 (.25)****	.43 (.18)***	.56 (.31)****
D	70	.45 (.20)****	.12 (.01)	.33 (.11)**
WS	74	.44 (.19)****	.28 (.08)*	.44 (.19)****
BR	68	.57 (.32)****	.26 (.07)*	.50 (.25)****
BM	66	.55 (.30)****	.40 (.16)***	.58 (.34)****
BWL	69	.54 (.29)****	.29 (.08)*	.50 (.25)****

Note. K-BIT = Kaufman Brief Intelligence Test; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; LWID = Letter-Word Identification; PC = Passage Comprehension; C = Calculation; AP = Applied Problems; D = Dictation; WS = Writing Samples; BR = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language. r^2's presented in parentheses.
*$p < .05$. **$p < .01$. ***$p < .001$. ****$p < .0001$.

Table 6

Pearson product-moment correlation coefficients between the WISC-III and WJ-R

	n	WISC-III VIQ	PIQ	FSIQ
WJ-R ACH				
LWID	67	.30 (.09)**	.08 (.01)	.21 (.04)
PC	67	.57 (.32)****	.39 (.15)***	.53 (.28)****
C	66	.42 (.18)***	.29 (.08)*	.38 (.14)***
AP	66	.65 (.42)****	.46 (.21)****	.61 (.37)****
D	70	.28 (.08)*	.11 (.01)	.21 (.04)
WS	74	.58 (.34)****	.44 (.19)****	.56 (.31)****
BR	68	.46 (.21)****	.23 (.05)*	.38 (.14)***
BM	66	.63 (.40)****	.44 (.19)***	.59 (.35)****
BWL	69	.52 (.27)****	.35 (.12)**	.47 (.22)****

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; VIQ = Verbal IQ; PIQ = Performance IQ; FSIQ = Full Scale IQ; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; LWID = Letter-Word Identification; PC = Passage Comprehension; C = Calculation; AP = Applied Problems; D = Dictation; WS = Writing Samples; BR = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language. r^2's presented in parentheses.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. $^{****}p < .0001$.

Table 7 presents descriptive statistics for the WJ-R ACH and examination of mean scores indicated that this sample of students previously identified as learning disabled as a group scored approximately 1 *SD* below the mean of the standardization sample. Because of the heterogeneous nature of groups of "learning disabled" students (learning disability may exist in any *one* of 7 areas), clinical group mean scores may appear larger than expected when compared to mean WJ-R ACH scores for the standardization sample. Subjects in the present study scored significantly higher on Passage Comprehension than on Letter-Word Identification $t(66) = 6.57$, $p < .0001$; and scored significantly higher on Applied Problems than Calculation $t(65) = 11.88$, $p < .0001$. They also scored significantly higher on Writing Samples than Dictation, $t(69) = 9.02$, $p < .0001$. Comparisons between global scales indicated that subjects obtained significantly higher Broad Reading scores than Broad Written Language scores, $t(66) = 6.45$, $p < .0001$ and significantly

Table 7

Descriptive statistics for WJ-R Tests of Achievement

WJ-R ACH	*n*	*M*	*SD*	Range
Letter-Word Identification	67	79.96	10.41	53 - 101
Passage Comprehension	67	87.06	12.22	48 - 131
Calculation	66	77.02	10.41	53 - 102
Applied Problems	66	91.12	11.93	60 - 132
Dictation	70	71.04	10.31	32 - 88
Writing Samples	74	85.82	16.81	32 - 126
Broad Reading	68	81.25	11.27	52 - 113
Broad Mathematics	66	81.29	11.64	54 - 119
Broad Written Language	69	75.30	10.73	31 - 97

Note. WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement.

higher Broad Mathematics scores than Broad Written Language scores, $t(64) = 5.16$, $p < .0001$; but did not differ in Broad Reading and Broad Mathematics $t(64) = 0.23$, *ns*. These data indicated that these subjects seemed to perform better on the more abstract or complex subtests (Passage Comprehension, Applied Problems, and Writing Samples) of the WJ-R ACH than on the less complex subtests (Letter-Word Identification, Calculation, and Dictation).

Tables 8 through 16 present the 2 X 2 diagnostic efficiency tables and statistics for each of the WJ-R ACH subtests and global achievement scores and show that κ coefficients ranged from .31 to 1.0 ($Mdn_{\kappa} = .65$). All κ coefficients were significant and indicated that the agreement of the presence or absence of severe achievement-ability discrepancies between the K-BIT and WJ-R ACH with the presence or absence of severe achievement-ability discrepancies between the WISC-III and WJ-R ACH were well beyond chance. In fact, 78% (7 of 9) of the κ coefficients were in the substantial or almost perfect agreement range (Everitt & Hay, 1992). One κ coefficient indicated perfect agreement between the K-BIT and WISC-III in identifying presence or absence of severe discrepancy for the Applied Problems subtest. Other indexes of diagnostic efficiency also yielded encouraging, positive results. Agreement between the K-BIT and WISC-III in identifying severe discrepancies was reflected in high levels of positive predictive power, negative predictive power, and overall correct classification. Positive predictive power referred to the proportion of subjects with severe achievement-ability discrepancies identified by the K-BIT who truly showed a severe achievement-ability discrepancies with the WISC-III. Negative predictive power was indicated by the proportion of subjects who did not show a severe achievement-ability discrepancies with the K-BIT who likewise did not show a severe achievement-ability discrepancies with the WISC-III. Overall correct classification is the proportion of subjects correctly classified with and without severe achievement-ability discrepancies with the K-BIT. Generally low false positive rates were also observed. False positive classifications were indicated by subjects who showed severe achievement-ability discrepancies by the K-BIT who did not show the corresponding severe achievement-ability discrepancies with the WISC-III. False negative rates were, in some cases, moderately high. False negative classifications were indicated by those who showed no severe achievement-ability discrepancies with the K-BIT but showed severe achievement-ability discrepancies with the WISC-III. In a reevaluation situation, a false positive classification would result in continuing to classify a student's achievement as "discrepant" when it "truly" is not (based on a comprehensive intellectual ability measure). The result would likely be continuing eligibility for special education for that student as when only using a brief intellectual measure. A false negative classification would result in determining that the student's achievement is not "discrepant" when it "truly" was discrepant (based on a comprehensive intellectual ability measure). This result would likely be the terminating of special education programming eligibility when only using a brief intellectual measure. Both of these situations are possible in any assessment using a brief or comprehensive intellectual measure due to measurement error and there is yet no agreement as to which is the more serious error (Heath & Kush, 1991; Reynolds, 1984).

Table 8

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Letter-Word Identification (LWID)

		WISC-III--LWID		
		SD	No SD	Total
	SD	7	2	9
K-BIT--LWID	No SD	5	53	58
	Total	12	55	67

Sensitivity	.58
Specificity	.96
Positive Predictive Power	.78
Negative Predictive Power	.91
False Positive Rate	.04
False Negative Rate	.42
Overall Correct Classification (Hit) Rate	.90
κ	.61
SE_{κ}	.12
z	5.03
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; LWID = Letter-Word Identification; SD = Severe Discrepancy.

Table 9

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Passage Comprehension (PC)

		WISC-III--PC		
		SD	No SD	Total
	SD	2	0	2
K-BIT--PC	No SD	2	63	65
	Total	4	63	67

Sensitivity	.50
Specificity	1.00
Positive Predictive Power	1.00
Negative Predictive Power	.97
False Positive Rate	.00
False Negative Rate	.50
Overall Correct Classification (Hit) Rate	.97
κ	.65
SE_{κ}	.11
z	5.71
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; PC = Passage Comprehension; SD = Severe Discrepancy.

Table 10

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Calculation (C)

		WISC-III--C		
		SD	No SD	Total
	SD	11	3	14
K-BIT--C	No SD	3	49	52
	Total	14	52	66

Sensitivity	.79
Specificity	.94
Positive Predictive Power	.79
Negative Predictive Power	.94
False Positive Rate	.06
False Negative Rate	.21
Overall Correct Classification (Hit) Rate	.91
κ	.73
SE_{κ}	.12
z	5.92
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; C = Calculation; SD = Severe Discrepancy.

Table 11

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Applied Problems (AP)

		WISC-III--AP		
		SD	No SD	Total
K-BIT--AP	SD	1	0	1
	No SD	0	65	65
	Total	1	65	66

Sensitivity	1.00
Specificity	1.00
Positive Predictive Power	1.00
Negative Predictive Power	1.00
False Positive Rate	.00
False Negative Rate	.00
Overall Correct Classification (Hit) Rate	1.00
κ	1.00
SE_κ	.13
z	7.58
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; AP = Applied Problems; SD = Severe Discrepancy.

Table 12

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Dictation (D)

		WISC-III--D		
		SD	No SD	Total
	SD	26	6	32
K-BIT--D	No SD	5	33	38
	Total	31	39	70

Sensitivity	.84
Specificity	.85
Positive Predictive Power	.81
Negative Predictive Power	.87
False Positive Rate	.15
False Negative Rate	.16
Overall Correct Classification (Hit) Rate	.84
κ	.68
SE_{κ}	.12
z	5.71
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; D = Dictation; SD = Severe Discrepancy.

Table 13

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Writing Samples (WS)

		WISC-III--WS		
		SD	No SD	Total
K-BIT--WS	SD	6	2	8
	No SD	3	63	66
	Total	9	65	74

Sensitivity	.67
Specificity	.97
Positive Predictive Power	.75
Negative Predictive Power	.95
False Positive Rate	.03
False Negative Rate	.33
Overall Correct Classification (Hit) Rate	.93
κ	.67
SE_{κ}	.12
z	5.75
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; WS = Writing Samples; SD = Severe Discrepancy.

Table 14

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Broad Reading (BR)

		WISC-III--BR		
		SD	No SD	Total
K-BIT--BR	SD	4	1	5
	No SD	6	57	63
	Total	10	58	68

Sensitivity	.40
Specificity	.98
Positive Predictive Power	.80
Negative Predictive Power	.90
False Positive Rate	.02
False Negative Rate	.60
Overall Correct Classification (Hit) Rate	.90
κ	.48
SE_{κ}	.11
z	4.29
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; BR = Broad Reading; SD = Severe Discrepancy.

Table 15

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Broad Mathematics (BM)

		WISC-III--BM		
		SD	No SD	Total
	SD	2	4	6
K-BIT--BM	No SD	3	57	60
	Total	5	61	66

Sensitivity	.40
Specificity	.93
Positive Predictive Power	.33
Negative Predictive Power	.95
False Positive Rate	.07
False Negative Rate	.60
Overall Correct Classification (Hit) Rate	.89
κ	.31
SE_{κ}	.12
z	2.50
$p <$	.01

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; BM = Broad Mathematics; SD = Severe Discrepancy.

Table 16

Frequencies of subjects showing severe discrepancies based on WISC-III Full Scale IQ and K-BIT IQ Composite estimated achievement for Broad Written Language (BWL)

		WISC-III--BWL		
		SD	No SD	Total
	SD	12	3	15
K-BIT--BWL	No SD	6	48	54
	Total	18	51	69

Sensitivity	.67
Specificity	.94
Positive Predictive Power	.80
Negative Predictive Power	.89
False Positive Rate	.06
False Negative Rate	.33
Overall Correct Classification (Hit) Rate	.87
κ	.64
SE_{κ}	.12
z	5.37
$p <$	.0001

Note. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; K-BIT = Kaufman Brief Intelligence Test; BWL = Broad Written Language; SD = Severe Discrepancy.

General Discussion

The K-BIT IQ Composite, Vocabulary, and Matrices scores compared favorably to the WISC-III IQ's and Index scores and these data provided ample evidence supporting the concurrent validity of the K-BIT as a reliable and valid brief estimate of *general* intellectual abilities in reevaluation of elementary and middle school children with learning disability. The comparison of Vocabulary and Matrices subtest scores, however, appears to be a questionable practice based on these data as it does not provide significant insight into possible verbal and nonverbal differences in a comprehensive intellectual measure. Kaufman and Kaufman (1990) were justifiably cautious in recommending that no inferences be made about possible verbal and nonverbal differences with the K-BIT, however, they provide no theoretical or empirical support for the "mandate" (p.46) for recommending administration of a comprehensive intellectual battery to investigate abnormal Vocabulary - Matrices discrepancies. The present study suggests that this "mandate" may not be justified given the low positive predictive power, low sensitivity, and low, nonsignificant κ coefficients. These results seem to fit theoretical conceptualizations of the instability of verbal-nonverbal (VIQ-PIQ) differences (Macmann & Barnett, 1994). Further research is obviously needed to address this issue.

More importantly, the present study found the K-BIT to be extremely useful in that achievement-ability discrepancies found between the K-BIT and WJ-R ACH had very high positive and negative predictive power as well as very high overall correct classification when compared to achievement-ability discrepancies found between the WISC-III and WJ-R ACH. If the present results are replicated, then the K-BIT may supplant a comprehensive intellectual ability measure in reevaluations of students with SLD while retaining a high degree of diagnostic agreement. This practice could save considerable time in the reevaluation process that could be spent in alternative assessment practices or in providing alternative services such as consultation, program development, counseling, or research.

Future research should continue to examine the relationship of the K-BIT with other comprehensive intellectual ability measures and with different samples of normal and clinical groups in order to further define and delineate it's psychometric characteristics. Differences between racial or ethnic groups as well as bilingual subjects should also be explored. As with comprehensive intellectual ability measures, it will be important to determine if there is differential validity for different subgroups in the population. Replication of the present findings in future research may establish the K-BIT as the standard for quickly estimating the general intelligence of individuals.

References

Anastasi, A. (1988). *Psychological Testing* (6th Ed.). New York: Macmillan.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Canivez, G. L. (1995). Validity of the Kaufman Brief Intelligence Test: Comparisons with the Wechsler Intelligence Scale for Children-Third Edition. *Assessment, 2,* 101-111.

Canivez, G. L., & Watkins, M. W. (in press). Automated calculation of diagnostic efficiency statistics. *Behavior Research Methods, Instruments, and Computers.*

Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.

Everitt, B. S., & Hay, D. F. (1992). *Talking about statistics: A psychologist's guide to data analysis.* New York: Halsted.

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions* (2nd Ed.). New York: Wiley.

Fleiss, J. L., Cohen, J., Everitt, B. S. (1969). Large sample standard errors of Kappa and weighted Kappa. *Psychological Bulletin, 72*, 323-327.

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th Ed.). New York: McGraw-Hill.

Goh, D. S., Teslow, C. J., & Fuller, G. B. (1981). The practice of psychological assessment among school psychologists. *Professional Psychology, 12*, 696-706.

Heath, C. P., & Kush, J. C. (1991). Use of discrepancy formulas in the assessment of learning disabilities. In J. E. Obrzut and G. W. Hynd (Eds.), *Neuropsychological foundations of learning disabilities: A handbook of issues, methods, and practice* (pp. 287-307). San Diego, CA: Academic Press.

Hutton, J. B., Dubes, R., & Muir, S. (1992). Assessment practices of school psychologists: Ten years later. *School Psychology Review, 21*, 271-284.

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

Kessel, J. B., & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires: Extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment, 5*, 395-399.

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.

Naglieri, J. A. (1993). Pairwise and ipsative comparisons of WISC-III IQ and Index scores. *Psychological Assessment, 5*, 113-116.

Naugle, R. I., Chelune, G. J., & Tucker, G. D. (1993). Validity of the Kaufman Brief Intelligence Test. *Psychological Assessment, 5*, 182-186.

Prewett, P. N. (1992a). The relationship between the Kaufman Brief Intelligence Test (K-BIT) and the WISC-R with referred students. *Psychology in the Schools, 29*, 25-27.

Prewett, P. N. (1992b). The relationship between the Kaufman Brief Intelligence Test (K-BIT) and the WISC-R with incarcerated juvenile delinquents. *Educational and Psychological Measurement, 52*, 997-982.

Reschly, D. J., Genshaft, J., & Binder, M. S. (1987). *The 1986 NASP survey: Comparison of practitioners, NASP leadership, and university faculty on key issues.* Washington, DC: National Association of School Psychologists.

Reynolds, C. R. (1984). Critical measurement issues in learning disabilities. *Journal of Special Education, 18*, 1-26.

Ross-Reynolds, J. (1990) Best practices in conducting reevaluations. In A. Thomas and J. Grimes (Eds.), *Best Practices in School Psychology-II* (pp. 195-206). Washington, DC: National Association of School Psychologists.

Smith, D. K. (1984). Practicing school psychologists: Their characteristics, activities, and populations served. *Professional Psychology: Research and Practice, 15*, 798-810.

Thorndike, R. L., & Hagen, E. P. (1977). *Measurement and evaluation in psychology and education* (4th Ed.). New York: Wiley.

Telzrow, C. F. (1990). Best practices in reducing error in identifying specific learning disabilities. In A. Thomas and J. Grimes (Eds.), *Best Practices in School Psychology-II* (pp. 607-620). Washington, DC: National Association of School Psychologists.

United States Department of Education (1992). Assistance to states for the education of children with disabilities program and preschool grants for children with disabilities. *Federal Register, 57*, 44794-44852.

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

Wilson, L. R., & Cone, T. (1984). The regression equation method of determining academic discrepancy. *Journal of School Psychology, 22*, 95-110.

Woodcock, R. W., & Mather, N. (1989). WJ-R Tests of Achievement: Examiner's manual. In R. W. Woodcock & M. B. Johnson, *Woodcock-Johnson Psycho-Educational Battery-Revised*. Allen, TX: DLM Teaching Resources.