

Validity of the General Conceptual Ability Score From the Differential Ability Scales as a Function of Significant and Rare Interfactor Variability

Kasey M. Kotz
Pennsylvania State University

Marley W. Watkins
Arizona State University

Paul A. McDermott
University of Pennsylvania

Abstract. Some researchers have argued that discrepant broad index scores invalidate IQs, but others have questioned the fundamental logic of that argument. To resolve this debate, the present study used a nationally representative sample of children ($N = 1,200$) who were matched individually for IQ. Children with significantly uneven broad index score profiles and those with even broad index score profiles had equivalent reading and math skills. Discrepant broad index scores found in only 15%, 10%, 5%, and 1% of the population, respectively, also failed to differentially predict academic achievement. In addition, significantly higher Verbal broad index scores were not differentially predictive of reading achievement, nor were significantly higher Nonverbal/Spatial broad index scores differentially predictive of math achievement. It was concluded that the global ability score is the most parsimonious predictor of academic achievement, despite the presence of significant and rare variability among broad index scores.

Every year, millions of assessments that include a measure of intelligence are conducted (Sattler, 2001). Although only a portion of the competencies included in human intelligence is assessed, the results of these ability measures provide the best available long-range predictors of student achievement, school adjustment, level of vocational attain-

ment, and job performance (Gottfredson, 1997, 2004; Kubiszyn et al., 2000; Sattler, 2001; Schmidt & Hunter, 2004). Therefore, the measurement of intelligence can aid in the prediction of a range of important educational and occupational criteria (Braden, 1997). Consequently, examiners ought to render accurate and relevant interpretations of intelligence measures.

This paper is based on the MS thesis of the first author, which was supervised by the second author. The authors are grateful to Professors Robert Hale and Barbara Schaefer for their guidance.

Correspondence regarding this article should be addressed to Kasey M. Kotz, Pennsylvania State University, Department of Educational and School Psychology and Special Education, 111 CEDAR Building, University Park, PA 16802; E-mail: kmk354@psu.edu

Copyright 2008 by the National Association of School Psychologists, ISSN 0279-6015

The omnibus or global Intelligence Quotient (IQ) is commonly used as a predictor of academic achievement. Many clinicians, however, attempt to extract additional information from lower level subcomponents of the intelligence test, such as broad index scores and subtest scores (Donders, 1996; Kaufman, 1994; Kaufman & Lichtenberger, 2000). These lower level scores are then organized to form profiles of broad index or subtest scores. Although analysis of subtest scores has been historically popular, subtests should not be used for the prediction of achievement (Glutting, McDermott, Konold, Snelbaker, & Watkins, 1998; McDermott & Glutting, 1997; Watkins & Glutting, 2000). Even advocates of subtest analysis admit that this method should not be used for more than the examination of individual differences within a particular client (Kaufman, 1994; Sattler, 2001).

Unlike subtest analysis, broad index score analysis has increased in popularity (Prifitera, Saklofske, & Weiss, 2005). Some psychologists have argued that interfactor discrepancies *invalidate* the IQ, especially for the prediction of achievement (Hale & Fiorello, 2001; Hale, Fiorello, Kavanagh, Hoepfner, & Gaither, 2001; Lichtenberger, Kaufman, & Lai, 2002). This position has gained wide currency in clinical practice (Pfeiffer, Reddy, Kletzel, Schmelzer, & Boyer, 2000), and devaluation of the IQ in the presence of an uneven broad index score profile is often suggested in the professional literature (Drummond, 2004; Kaufman, 1994; Kaufman & Lichtenberger, 2002; Lezak, 1995; Sattler & Dumont, 2004; Weiss, Saklofske, & Prifitera, 2005). However, some researchers have challenged this interpretation and suggested that the IQ is the best predictor of academic outcomes despite broad index score variability (Glutting, Youngstrom, Ward, Ward, & Hale, 1997; Kline, Snyder, Guilmette, & Castellanos, 1993; Oh, Glutting, Watkins, Youngstrom, & McDermott, 2004; Watkins & Kush, 1994; Youngstrom, Kogos, & Glutting, 1999). In fact, Glutting, Watkins, Konold, and McDermott (2006) asserted that when observed ability scores are used to predict achievement,

only the IQ, an indicator of general intelligence, is needed.

Thus, two sets of researchers have analyzed similar data sets and arrived at contradictory conclusions. Glutting et al. (1997) suggested that the “global ability score will generally prove to be the most useful in predicting concurrent achievement” (p. 300), whereas Hale and Fiorello (2001) encouraged practitioners to “never interpret the global IQ score if there is significant scatter or score variability” (p. 132). Although both research groups used multiple regression, one group used hierarchical regression to identify the most parsimonious predictor, and the other group applied commonality analysis to specify an explanatory model. Given these conflicting recommendations, clinicians do not know how to interpret IQs in the presence of an uneven broad index score profile: Is the IQ invalid (as per Hale & Fiorello, 2001) or is it preferred over the variable broad index scores (as per Glutting et al., 1997)?

An alternative research design can be used to resolve this debate. Specifically, a matched case design can be used in which two groups are matched according to IQs and demographic characteristics. One group would be comprised of cases with a significantly uneven broad index score profile, and the other group would consist of matched cases with little broad index score variability. These two groups would then be compared to discover whether attendant predictions of academic achievement were equally accurate. This method of analysis would eliminate the problem of collinear variables in multiple regressions, yet it would allow prediction of academic achievement. In this manner, a matched case design can determine the predictive validity of IQ in the presence of a significantly uneven broad index score profile. In addition, the use of alternative methods to determine validity adheres to the most recent version of the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), which requires that validity be estab-

lished by examining several different processes of validation.

Following similar logic, a matched case design was used to examine the predictive validity of the Wechsler Adult Intelligence Scale—Third Edition (WAIS-III; Wechsler, 1997a), in which half of the participants had high intersubtest scatter and the other half had low intersubtest scatter (Ryan, Kreiner, & Burton, 2002). Multiple regression was used to determine whether the amount of intersubtest scatter influenced the relationship between IQ and memory. Predictions of Wechsler Memory Scale—Third Edition (Wechsler, 1997b) indices did not differ significantly between the groups of participants with low and high WAIS-III intersubtest scatter. Based on these results, Ryan et al. (2002) concluded that intersubtest scatter did not reduce the validity of the Full Scale Intelligence Quotient in predicting memory performance.

Although marked WAIS-III intersubtest scatter did not impair the ability of the WAIS-III IQ to predict memory outcomes, further research is needed. First, results based on the Wechsler Memory Scale—Third Edition may not generalize to academic achievement (Ryan et al., 2002). Also, this study employed the WAIS-III as the cognitive measure, but other ability assessment instruments have yet to be examined. Finally, the participants in this study were adult medical center patients. Thus, the results do not provide guidance for the interpretation of IQs obtained by school-age children.

The validity of the IQ as a predictor of academic achievement when broad index scores are discrepant remains unclear. Because the Differential Ability Scales (DAS; Elliott, 1990a) have been shown to produce reliable and valid indicators of cognitive ability (Aylward, 1992; Reinehr, 1992), the present study investigated the effect of a statistically significant, uneven broad index score profile on the predictive validity of the DAS General Conceptual Ability (GCA). In addition to statistical significance, the present study also considered clinical significance. If the results indicate that GCAs in the presence of statistically significant discrepancies be-

tween broad index scores do not predict academic achievement as well as those with non-discrepant broad index scores, no further examination would be needed. On the other hand, if the results suggest that statistically significant discrepancies between broad index scores do not differentially predict achievement, clinical significance would then be examined. The present study thus inspected those broad index scores that were statistically significant *and* rare in the population. This examination began with a 15% base rate, then successively examined 10%, 5%, and 1% base rates, respectively (Dori & Gordon, 2004). The use of these subsamples, which include cases with larger broad index score discrepancies than in the parent sample, will allow for a closer examination of the effects of an uneven broad index score profile. Such results would enable examiners to understand the difference, if any, between broad index scores needed to render the GCA invalid for the prediction of academic achievement.

Furthermore, the proposed study also acknowledges the possibility that the direction of the discrepancy may influence the validity of the GCA in predicting respective achievement scores. Therefore, a group with only significantly higher Verbal broad index scores was matched with nondiscrepant broad index scores, and the differences in prediction of academic achievement between groups were examined. Similarly, a group with only significantly higher Nonverbal or Spatial broad index scores was matched with nondiscrepant broad index scores, and the differences in prediction of academic achievement were measured. Finally, each of these groups was reconstituted to include subsamples at the 15%, 10%, and 5% prevalence levels to examine the level of discrepancy needed to invalidate the GCA for the High Verbal and High Nonverbal/Spatial groups.

Method

Participants

Participants were 1,200 children who completed the DAS during the national stan-

standardization of the Adjustment Scales for Children and Adolescents (McDermott, Marston, & Stott, 1993). The sample represented the population of school-age children ages 6 years, 0 months through 17 years, 11 months living in the United States in 1993. The participants were recruited from 201 school systems located in 70 U.S. Census statistical areas (metropolitan, suburban, rural) across four regions of the country (Northeast, Midwest, South, West). Within each age level, the sample conformed to parameters of the 1990 U.S. Census (U.S. Department of Commerce, 1990) for the variables of gender, ethnicity, parents' educational attainment, and geographic region.

Measures

Cognitive ability. Although the second edition of the DAS (DAS-II; Elliott, 2006) was recently published, the DAS in its original form was used in the current study for several reasons. First, DAS-II administration and scoring materials continue to be revised (see <http://harcourtassessment.com> for details of March 2007 corrections). Secondly, conormed achievement measures were not developed with the DAS-II as they were for the DAS. Finally, the DAS and the DAS-II contain 17 of the same subtests. Therefore, the DAS was used to determine general level of cognitive ability and academic achievement for the participants in this study. Of course, the DAS-II would be more appropriate for clinical practice given its modern norms (American Psychological Association, 2002).

The general indicator score obtained from the DAS is the GCA. For school-age children (6 years, 0 months through 17 years, 11 months), there are three indicator scores reflecting individual abilities: the Verbal Ability broad index score, the Nonverbal Reasoning broad index score, and the Spatial Ability broad index score. Each of the aforementioned broad indices is comprised of two core subtests. The Verbal Ability broad index measures complex verbal mental processing, with each subtest requiring verbal responses. The Nonverbal Reasoning broad index is a

measurement of nonverbal inductive reasoning that requires some complex mental processing skills. The Spatial Ability broad index is a measure of complex visual-spatial processing and requires the abilities to perceive and remember spatial relationships and shapes (Elliott, 1990b).

The reliability of the GCA is high at all ages, ranging from .90 to .95. Internal validity was established by using both confirmatory and exploratory factor analysis at various age levels. Criterion-related validity was assessed via correlations with other cognitive measures. Correlations of the DAS with the Stanford-Binet Intelligence Scale—Fourth Edition (Thorndike, Hagen, & Sattler, 1986) indicated similar mean scores for the composite ($r = .77$). Additional high correlations (ranging from .84 to .91) were found between the composite scores of the DAS and the Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974) at all age levels. The DAS has been shown to yield nonsignificantly different IQs and broad index scores from other measures, such as the WISC-III, for both White and minority children (DiCerbo & Barona, 2000). The DAS Verbal Ability broad index was highly correlated with such measures of verbal ability as the Verbal IQ of the Wechsler Intelligence Scale for Children—Revised ($r = .84$) and the verbal factor of the Stanford-Binet Intelligence Scale—Fourth Edition ($r = .72$). The DAS Nonverbal Reasoning broad index demonstrated strong relationships with other similar measures, including the Stanford-Binet Intelligence Scale—Fourth Edition Abstract-Visual Reasoning factor ($r = .76$) and Quantitative Reasoning factor ($r = .75$). The Spatial Ability broad index also correlated highly with other measures, such as the Wechsler Intelligence Scale for Children—Revised Performance IQ ($r = .77$) and the Abstract-Visual Reasoning Factor of the Stanford-Binet Intelligence Scale—Fourth Edition ($r = .67$). The DAS has also been found valid for distinguishing between students who are at risk and those who are within the normal range of school functioning (McIntosh, 1999).

Achievement criteria. The DAS contains two individual achievement scales that were conormed with the cognitive ability tests: Word Reading ($\alpha = .92$) and Basic Number Skills ($\alpha = .87$).¹ Word Reading examines children's ability to recognize and orally read single words. Basic Number Skills is a test of children's knowledge of the concepts and skills that underlie basic competence in arithmetic calculation.

Each of these tests demonstrates moderate to high positive correlations with other individually administered achievement tests ($r = .64-.88$) and group-administered achievement tests ($r = .62-.77$; Elliott, 1990b). In addition, these achievement tests show moderately good convergent validity with school performance, specifically teacher-assigned grades ($r = .43-.60$; Elliott, 1990b). These coefficients are similar to those obtained by other commonly used achievement tests, including the Wechsler Individual Achievement Test—Second Edition (Wechsler, 2001).

Procedure

Two groups were constituted so that they differed on degree of broad index score discrepancy on the DAS, but matched on GCA. The group with large broad index score discrepancies was labeled the *uneven profile* group and the group with small broad index score discrepancies was named the *even profile* group. To create these groups, the DAS broad index scores of all 1,200 participants were scrutinized for significant variability. The average significant difference between broad index scores at the $p = .05$ level is 16.8 points for the Verbal and Nonverbal Reasoning scores, 15.9 points for the Verbal and Spatial scores, and 15.3 points for the Nonverbal Reasoning and Spatial scores (Elliott, 1990b). Thus, a broad index score discrepancy of 17 or greater was used to determine the uneven profile group for cases with differences between Verbal and Nonverbal Reasoning scores, and a discrepancy of 16 or more was used to identify the uneven profile group for cases with differences between Verbal and

Spatial and between Nonverbal Reasoning and Spatial scores. These specific discrepancies were employed to emulate standard clinical practice and followed from Elliott (2005).

The resulting uneven profile group contained 503 participants. Thus, 42% of the normative sample exhibited at least one significant broad index discrepancy. A student with nonsignificant broad index discrepancies was then individually matched on GCA score (± 1 point) to each student in the uneven profile group. Following the GCA match, each student was matched on the closest age in years. Further matching was sequentially performed for sex, ethnicity, years of parent education, and region of the country. Unfortunately, some cases could not be exactly matched on all demographic variables (see Tables 1 and 2 in the Results section).

Following clinical practice (Sattler & Dumont, 2004) and Table B.3 in Elliott (1990b), subgroups were identified based on rarity or prevalence of broad index score differences. The first uneven profile subgroup was comprised of only those cases that had a 15% base rate or prevalence, which translated to cases with discrepancies $\geq 19-21$ points ($n = 381$). Following this, base rates of 10% (discrepancies $\geq 21-24$ points, $n = 275$), 5% (discrepancies $\geq 25-28$ points, $n = 165$), and 1% (discrepancies $\geq 34-37$ points, $n = 28$) were used to form new uneven profile subgroups, which were all matched with their respective even profile cases from the parent profile groups. Thus, the case membership of the rare prevalence subgroups was nested within the parent profile groups.

¹A Spelling achievement scale ($\alpha = .92$) was also conormed with the ability tests, which focuses on children's ability to produce correct spellings of phonetically regular and irregular words. However, spelling is not recognized in current special education law (Individuals with Disabilities in Education Improvement Act, 2004) and is only one of several lower level skills within written expression (Vaughn, Gersten, & Chard, 2000). Consequently, spelling achievement was excluded from subsequent analyses.

Because it could be argued that the direction of the broad index score discrepancy determines the predictive validity of the global intelligence score, several supplemental analyses were undertaken (see Tables 4–7 in the Results section for descriptive statistics). The first analysis examined whether a significantly higher Verbal broad index score contributed to differential prediction of achievement. Those cases with Verbal broad index scores significantly higher than either the Spatial or Nonverbal broad index scores were used. The respective matched even profile group cases were also selected for this analysis.

Relatedly, an additional analysis examined significantly higher Nonverbal broad index scores to determine if they rendered the GCA differentially predictive of achievement. Those cases in the parent uneven profile group with Nonverbal or Spatial broad index scores significantly higher than the Verbal broad index score were used. Their respectively matched even profile group cases were also included in this analysis. The uneven profiles for the Nonverbal and Spatial broad indexes were combined because these are measures of inductive reasoning and visual-spatial processing, which are both necessary to complete math problems (Elliott, 1990b).

Finally, these high Verbal and high Nonverbal/Spatial matched subgroups were further subdivided to form additional subgroups. These nested groups only included cases in which the broad index score discrepancies were both significant and rare at the 15%, 10%, and 5% prevalence levels. There were too few cases for analysis at the 1% prevalence level.

Analyses

It was hypothesized that the relationship between GCA (independent variable) and achievement (dependent variable) will differ depending on the level of broad index score variability (moderator variable). Similar moderator effects are involved when predictive test bias is hypothesized (Reynolds, 1995). For example, the relationship between IQ (independent variable) and achievement (dependent variable) will differ depending on the ethnicity

(moderator variable) of the student. Moderated multiple regression is recommended for testing such hypotheses (Society for Industrial and Organizational Psychology, 2003; Stone-Romero & Anderson, 1994). Using this method, reading and mathematics subtest scores were sequentially regressed on the GCA, profile group membership (a dichotomous categorical indicator), and an interaction term between the GCA and profile group.

Centering was used for each of the regression analyses to control for multicollinearity (Kraemer & Blasey, 2004). In addition, because the intercept is defined as the value of the criterion when each predictor equals zero, centering ensures that the intercept is the value of the criterion when the predictor is equal to the mean value. Therefore, in the current study, the intercept is equal to the achievement score when the GCA is at the mean value. In contrast, the slope indicates the increase in achievement score per one point increase in GCA. A statistically significant main effect for profile group membership or interaction between GCA and profile group membership would signal differential predictive validity. Because multiple significance tests were conducted (i.e., 78) with overlapping participants, a conservative alpha level ($p < .01$) was adopted for each test to partially control the overall Type I error rate.

Results

Descriptive statistics for the cognitive and academic achievement measures as well as age, parent education, ethnicity, sex, and regional characteristics of participants are presented in Tables 1 and 2. IQ and age differences between even and uneven profile groups were examined with t tests and did not differ significantly ($p < .05$). Standardized mean differences ranged from .00 to .07, well within Cohen's (1988) *small* effect size category. Differences in parent education, ethnicity, sex, and region were analyzed with χ^2 tests. As expected, the even and uneven profile groups did not differ significantly on any of the five demographic variables ($p < .05$), thus matching, although not perfectly, created groups that

Table 1
General Conceptual Ability, Age, and Academic Achievement for Broad Index Score Variability Groups Across Prevalence Levels

Variable	Even Profile Group				Uneven Profile Group				<i>d</i>
	<i>M</i>	<i>SD</i>	Skew	<i>SE</i>	<i>M</i>	<i>SD</i>	Skew	<i>SE</i>	
Statistically Significant Broad Index Score Variability (<i>n</i> = 1,006)									
GCA	100.66	13.74	.02	.11	100.66	13.74	.02	.11	0.00
Age (years)	11.27	3.35	.06	.11	11.51	3.31	-.02	.11	0.07
Basic Number Skills	100.33	13.31	.17	.11	100.91	15.10	.06	.11	0.04
Word Reading	101.53	13.59	.08	.11	100.21	15.18	.02	.11	0.09
15% Base Rate (<i>n</i> = 762)									
GCA	100.80	13.79	.03	.13	100.80	13.78	.03	.13	0.00
Age (years)	11.46	3.33	-.01	.13	11.66	3.29	-.06	.13	0.06
Basic Number Skills	100.20	13.43	.24	.13	100.99	14.98	.12	.13	0.06
Word Reading	101.70	13.82	.05	.13	100.45	15.59	.04	.13	0.09
10% Base Rate (<i>n</i> = 550)									
GCA	101.46	14.01	-.02	.15	101.46	14.01	-.02	.15	0.00
Age (years)	11.65	3.26	-.03	.15	11.80	3.28	-.09	.15	0.05
Basic Number Skills	100.41	13.20	.27	.15	101.56	15.28	.03	.15	0.08
Word Reading	101.71	13.03	.07	.15	100.95	15.74	.10	.15	0.05
5% Base Rate (<i>n</i> = 330)									
GCA	102.25	13.41	-.07	.19	102.25	13.41	-.07	.19	0.00
Age (years)	11.98	3.21	-.07	.19	11.98	3.27	-.16	.19	0.00
Basic Number Skills	101.39	13.78	.08	.19	101.48	15.49	-.07	.19	0.01
Word Reading	102.72	13.93	-.02	.19	101.60	16.32	.17	.19	0.07
1% Base Rate (<i>n</i> = 56)									
GCA	103.64	12.91	-.20	.44	103.64	12.91	-.20	.44	0.00
Age (years)	11.43	3.21	.10	.44	11.32	3.54	.14	.44	0.03
Basic Number Skills	103.11	14.00	-.12	.44	100.82	13.96	.55	.44	0.00
Word Reading	102.61	12.17	-.23	.44	103.07	17.72	.15	.44	0.03

Note. GCA = General Conceptual Ability. Means and standard deviations of GCA, age, Basic Number Skills, and Word Reading were not significantly ($p < .05$) different between even and uneven profile groups at any base rate. *d* refers to Cohen's statistic for effect size.

were relatively equivalent on GCA and demographic characteristics.

For both groups the GCA was moderately correlated with both the Basic Number Skills subtest ($r_{\text{even profile}} = .52$, $r_{\text{uneven profile}} = .57$) and the Word Reading subtest ($r_{\text{even profile}} = .53$, $r_{\text{uneven profile}} = .57$). These relationships are very close to those observed in the DAS school-age standardization sample ($r = .60$). Observed standard deviations of the GCA were also generally close to the expected value

of 15, suggesting that restriction of range did not likely attenuate correlations for either the even or uneven profile groups.

As shown in Table 3, regression results indicated no significant incremental prediction of Word Reading by profile group or the interaction term. Similarly, there was no significant incremental prediction of Basic Number Skills by profile group or the interaction term. Therefore, the GCAs of both the even and uneven profile groups were equally predictive

Table 2
Frequencies for Broad Index Score Variability Groups by Sex, Ethnicity,
National Region, and Parent Education Level

Variable	Statistically Significant Group		15% Base Rate Group		10% Base Rate Group		5% Base Rate Group		1% Base Rate Group	
	Even	Uneven	Even	Uneven	Even	Uneven	Even	Uneven	Even	Uneven
Sex										
Male	246	268	182	201	131	140	78	83	12	16
Female	257	235	199	180	144	135	87	82	16	12
Ethnicity										
Black	79	71	63	55	45	29	28	19	5	5
Hispanic	59	61	47	47	34	42	21	15	3	5
White	353	351	262	264	189	192	110	112	19	16
Other	12	20	9	15	7	12	6	8	1	2
Region										
North Central	128	128	99	94	68	60	41	32	10	7
Northeast	99	87	70	62	57	49	24	28	4	4
South	172	174	133	133	96	99	61	63	5	12
West	104	114	79	92	54	70	39	42	9	5
Parent education										
<12 years	97	79	74	60	51	47	30	28	4	6
High school	193	183	149	140	103	91	65	53	10	6
Some college	122	144	91	109	73	81	42	51	10	10
College degree	91	97	67	72	48	56	28	33	4	6

Note. There were 1,006 participants in the statistically significant broad index score variability group, 762 participants in the 15% base rate group, 550 participants in the 10% base rate group, 330 participants in the 5% base rate group, and 56 participants in the 1% base rate group. Even and uneven groups were not significantly ($p < .05$) different at any base rate in terms of sex, ethnicity, national region, and parent education.

of achievement in Basic Number Skills and Word Reading.

Moderated multiple regression was also used to analyze the prediction of both Basic Number Skills and Word Reading for the nested subgroups with 15%, 10%, 5%, and 1% prevalence levels. Again, no significant incremental prediction for Word Reading was found for rare profile groups or for the interaction terms. In addition, no significant incremental prediction for Basic Number Skills was found for rare profile groups or for the interaction terms. Thus, even large and rare broad index score differences did not render the GCA a statistically biased predictor of academic achievement.

To understand whether the direction of the discrepancy contributed to the validity of the GCA, moderated multiple regression was used to analyze the high Verbal uneven profile group and the high Nonverbal/Spatial uneven profile group with their respective matched groups. Descriptive statistics for the matched Verbal even and uneven profile groups are displayed in Tables 4 and 5, and the descriptive statistics for the high Nonverbal/Spatial uneven profile group and its matched cases are shown in Tables 6 and 7. As shown in Table 8, in the prediction of achievement on the DAS, all but one of the analyses for the high Nonverbal/Spatial group yielded nonsignificant predictions.

Table 3
Summary of Regression Results for Each Academic Achievement Area
Across Prevalence Levels

Test	Basic Number Skills			Word Reading		
	<i>R</i>	<i>R</i> ²	ΔR^2	<i>R</i>	<i>R</i> ²	ΔR^2
Statistically Significant Broad Index Score Variability (<i>n</i> = 1,006)						
GCA	.55	.30	.30*	.55	.30	.30*
+ Profile Group	.55	.30	.00	.55	.30	.00
+ GCA \times Group	.55	.30	.00	.55	.31	.01
15% Base Rate (<i>n</i> = 762)						
GCA	.54	.29	.29*	.53	.28	.28*
+ Profile Group	.54	.29	.00	.54	.29	.01
+ GCA \times Group	.54	.30	.01	.54	.29	.00
10% Base Rate (<i>n</i> = 550)						
GCA	.54	.30	.30*	.53	.28	.28*
+ Profile Group	.55	.30	.00	.53	.28	.00
+ GCA \times Group	.55	.30	.00	.54	.29	.01
5% Base Rate (<i>n</i> = 330)						
GCA	.52	.27	.27*	.54	.29	.29*
+ Profile Group	.52	.27	.00	.54	.30	.01
+ GCA \times Group	.52	.27	.00	.54	.30	.00
1% Base Rate (<i>n</i> = 56)						
GCA	.55	.30	.30*	.58	.34	.34*
+ Profile Group	.55	.31	.01	.88	.34	.00
+ GCA \times Group	.56	.31	.00	.63	.39	.05

Note. GCA = General Conceptual Ability.
 * $p < .001$.

Although one significant result was found for the prediction of Word Reading by profile group ($\Delta R^2 = .02$, $p = .001$), this increment has little practical significance and is small by Cohen's (1988) guidelines. In the prediction of achievement for the high Verbal profile group, Table 9 shows that no significant incremental prediction was found for the profile group nor for the interaction term.

Finally, after using moderated multiple regression to investigate these directional groups at the lower prevalence levels, no significant increment in prediction of reading achievement was found for profile group nor for the interaction term. Similarly, no sig-

nificant incremental prediction of math achievement was found for profile group or for the interaction term. Thus, despite the consideration of statistical significance, rarity, and direction of difference between broad index scores, the IQ remained a statistically unbiased predictor of academic achievement.

Discussion

Historically, intelligence tests were devised by Binet to measure students' ability to succeed in school. This fundamental characteristic of IQ tests has been empirically supported for more than 100 years (Kamphaus,

Table 4
DAS Cognitive Scores, Academic Achievement Scores, and Age for Even Profile and High Verbal Uneven Profile Groups Across Prevalence Levels

Variable	Even Profile Group		High Verbal Uneven Group		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Statistically Significant Broad Index Score Variability (<i>n</i> = 382)					
GCA	100.16	13.46	100.17	13.46	0.00
Age (years)	11.32	3.38	11.52	3.28	0.06
Verbal Broad Index*	100.09	11.97	111.70	12.40	0.95
Nonverbal Broad Index	100.05	12.98	95.93	14.66	0.30
Spatial Broad Index*	100.24	11.52	92.73	12.87	0.62
Word Reading	102.32	13.48	104.18	14.60	0.13
Basic Number Skills	99.90	12.51	101.11	14.28	0.09
15% Base Rate (<i>n</i> = 276)					
GCA	100.13	13.41	100.14	13.40	0.00
Age (years)	11.52	3.37	11.61	3.33	0.03
Verbal Broad Index*	99.74	11.93	113.02	12.26	1.10
Nonverbal Broad Index*	100.25	12.72	94.59	14.96	0.41
Spatial Broad Index*	100.34	11.73	92.70	13.04	0.62
Word Reading	102.41	13.79	105.07	14.18	0.19
Basic Number Skills	99.51	12.65	100.72	13.42	0.09
10% Base Rate (<i>n</i> = 188)					
GCA	102.06	13.13	102.06	13.13	0.00
Age (years)	11.38	3.37	11.61	3.33	0.04
Verbal Broad Index*	101.37	11.40	115.81	11.60	1.26
Nonverbal Broad Index	102.02	12.65	96.17	15.50	0.41
Spatial Broad Index*	101.74	11.67	93.14	13.32	0.69
Word Reading	104.18	13.00	107.11	14.00	0.22
Basic Number Skills	101.33	12.45	101.80	13.94	0.04
5% Base Rate (<i>n</i> = 114)					
GCA	104.25	11.22	104.25	11.22	0.00
Age (years)	11.54	3.19	11.54	3.25	0.00
Verbal Broad Index*	103.60	9.62	118.72	9.97	1.54
Nonverbal Broad Index	104.63	11.06	98.51	14.72	0.47
Spatial Broad Index*	102.51	10.62	93.39	13.40	0.75
Word Reading	106.74	12.40	109.09	14.59	0.17
Basic Number Skills	103.16	12.32	103.05	13.45	0.01

Note. GCA = General Conceptual Ability. *d* refers to Cohen's statistic for effect size.

**p* < .001.

2001). Subsequently, IQ has been found to be a powerful predictor of success in occupational training, vocational attainment, and life outcomes as diverse as accident prevention and susceptibility to health risks (Gott-

fredson, 1997, 2004; Schmidt & Hunter, 2004).

By necessity, IQ test batteries include subtest scores and first-order broad index scores in addition to the omnibus IQ. The

Table 5
Frequencies for the Even Profile and the High Verbal Uneven Profile
Groups by Sex, Ethnicity, National Region, and Parent Education Level

Variable	Statistically Significant Group		15% Base Rate Group		10% Base Rate Group		5% Base Rate Group	
	Even	Uneven	Even	Uneven	Even	Uneven	Even	Uneven
Sex								
Male	91	97	66	69	43	49	26	28
Female	100	94	72	69	51	45	31	29
Ethnicity								
Black	29	33	23	27	17	12	11	9
Hispanic	24	10	19	6	11	5	6	4
White	131	140	91	100	61	74	38	41
Other	7	8	5	5	5	3	2	3
Region								
North Central	38	55	28	42	21	27	13	10
Northeast	40	43	29	30	20	22	10	15
South	72	64	51	43	33	29	19	21
West	41	29	30	23	20	16	15	11
Parent education								
<12 years	36	20	26	13	12	10	4	7
High school	69	65	55	46	37	24	25	14
Some college	52	65	35	49	30	35	19	22
College degree	34	41	22	30	15	25	9	14

Note. There were 382 participants in the statistically significant broad index score variability group, 276 participants in the 15% base rate group, 188 participants in the 10% base rate group, and 114 participants in the 5% base rate group. Even and uneven groups were not significantly ($p < .01$) different at any base rate in terms of sex, ethnicity, national region, and parent education.

structure of intelligence into general and specific factors and the relationship of general and specific factors to external criteria are properly investigated with factor analysis and structural equation modeling (Carroll, 1993; Lubinski, 2000). These nomothetic multivariate statistical techniques attempt to reduce the effects of measurement error to better understand relationships among latent constructs. For example, factor analysis was used to delineate a three-tier hierarchical structure of intelligence (Carroll, 1993) and structural equation modeling was used to explicate the incremental influence of verbal skills beyond general intelligence in explaining academic achievement (Glutting et al., 2006).

However, clinicians work with manifest variables. In clinical practice, the omnibus IQ from an intelligence battery serves as a proxy for general intelligence, first-order broad index scores are assumed to be measures of specific abilities, and subtest scores are treated as narrow, third-order abilities. Of course, subtest scores are themselves formed by the accumulation of many items. Given these score levels, it is understandable that clinicians could choose to focus their interpretations on the very many items, many subtest scores, few broad index scores, or unitary IQ. Interpretation of items is rare given their poor reliability and weak validity. Although once popular, clinical

Table 6
DAS Cognitive Scores, Academic Achievement Scores, and Age for Even Profile and High Nonverbal/Spatial Profile Groups Across Prevalence Levels

Variable	Even Profile Group		Uneven Profile Group		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Statistically Significant Broad Index Score Variability (<i>n</i> = 456)					
GCA	101.07	14.32	101.06	14.31	0.00
Age (years)	11.36	3.30	11.61	3.27	0.08
Verbal Broad Index*	100.82	12.57	89.71	12.14	0.90
Nonverbal Broad Index*	100.57	13.13	105.11	15.59	0.32
Spatial Broad Index*	101.00	12.73	108.08	14.37	0.52
Basic Number Skills	100.11	14.22	100.86	15.87	0.05
Word Reading	100.73	13.73	97.07	15.36	0.25
15% Base Rate (<i>n</i> = 332)					
GCA	102.36	14.36	102.34	14.34	0.00
Age (years)	11.40	3.25	11.56	3.32	0.05
Verbal Broad Index*	101.97	12.68	89.57	12.14	1.00
Nonverbal Broad Index*	101.54	13.26	107.47	15.69	0.41
Spatial Broad Index*	102.17	12.61	109.12	14.59	0.51
Basic Number Skills	100.20	14.32	102.13	16.21	0.13
Word Reading	101.19	13.89	97.92	16.20	0.22
10% Base Rate (<i>n</i> = 246)					
GCA	102.31	14.05	102.28	14.04	0.00
Age (years)	11.38	3.33	11.50	3.28	0.04
Verbal Broad Index*	102.16	12.40	88.63	12.06	1.11
Nonverbal Broad Index*	101.22	12.98	108.27	15.50	0.49
Spatial Broad Index*	102.18	12.49	109.21	14.92	0.51
Basic Number Skills	99.69	14.17	102.46	15.34	0.19
Word Reading	100.03	13.56	99.54	14.24	0.04
5% Base Rate (<i>n</i> = 130)					
GCA	102.23	13.88	102.23	13.84	0.00
Age (years)	12.25	3.40	11.94	3.41	0.09
Verbal Broad Index*	101.89	12.23	86.49	12.22	1.26
Nonverbal Broad Index*	101.51	12.70	109.17	16.02	0.53
Spatial Broad Index*	102.05	12.65	110.58	14.22	0.63
Basic Number Skills	99.62	14.10	102.91	15.44	0.22
Word Reading	97.65	14.45	96.57	15.45	0.07

Note. DAS = Differential Ability Scales; GCA = General Conceptual Ability. *d* refers to Cohen's statistic for effect size. **p* < .001.

analysis of subtest scores has fundered from weak psychometric properties and poor diagnostic utility (Watkins, Glutting, & Youngstrom, 2005). Popular texts now recommend a focus on broad index scores, especially when there is significant interfactor

variability (American Psychiatric Association, 2000; Drummond, 2004; Elliott, 2005; Kaufman & Lichtenberger, 2000, 2002; Sattler & Dumont, 2004; Weiss et al., 2005).

Some researchers have suggested that significant interfactor variability *invalidates*

Table 7
Frequencies for the Even Profile and the High Nonverbal/Spatial Uneven Profile Groups by Sex, Ethnicity, National Region, and Parent Education Level

Variable	Statistically Significant Group		15% Base Rate Group		10% Base Rate Group		5% Base Rate Group	
	Even	Uneven	Even	Uneven	Even	Uneven	Even	Uneven
Sex								
Male	110	126	78	91	43	49	36	37
Female	118	102	88	75	51	45	29	28
Ethnicity								
Black	33	21	27	15	17	12	8	7
Hispanic	26	40	18	29	11	5	11	16
White	165	157	119	113	61	74	45	37
Other	4	10	2	9	5	3	1	5
Region								
North Central	66	51	50	33	21	27	18	9
Northeast	45	35	28	24	20	22	10	9
South	74	79	56	60	33	29	22	27
West	43	63	32	49	20	16	15	20
Parent education								
<12 years	51	42	37	30	12	10	16	13
High school	81	88	57	64	37	24	22	24
Some college	54	58	39	44	30	35	17	16
College degree	42	40	33	28	15	25	10	12

Note. There were 456 participants in the statistically significant broad index score variability group, 332 participants in the 15% base rate group, 188 participants in the 10% base rate group, and 130 participants in the 5% base rate group. Even and uneven groups were not significantly ($p < .01$) different at any base rate in terms of sex, ethnicity, national region, and parent education.

the global IQ (Hale & Fiorello, 2001). That claim was directly tested in this study and found to be false. Using a matched-case design, the present study compared the predictive validity of the DAS GCA between children with even and uneven broad index score profiles. No significant differences in the prediction of academic achievement were found. When the uneven profile group included only cases with interfactor discrepancies that were both significant and rare in the population at 15%, 10%, 5%, and 1% prevalence levels, there were still no significant differences found in the prediction of academic achievement. Furthermore, significantly higher Verbal broad index scores were not differentially predictive of achievement, nor were significantly

higher Nonverbal broad index scores differentially predictive of achievement. Thus, the GCA was a statistically unbiased predictor of academic achievement despite significant, rare, and directionally specific broad index score discrepancies.

The findings of the current study are in direct contradiction to the recommendations of the test developers, which advise test users to interpret profiles of strengths and weaknesses instead of the IQ. Furthermore, these results are also in opposition to the *Diagnostic and Statistical Manual of Mental Disorders* (APA, 2000), which advises the following:

When there is significant scatter in the subtest scores, the profile of relative strengths and weaknesses, rather than the

Table 8
Summary of Regression Results for Nonverbal/Spatial Broad Index Score to Each Academic Area Across Prevalence Levels

Test	Nonverbal/Spatial Broad Index to Basic Number Skills			Nonverbal/Spatial Broad Index to Word Reading		
	<i>R</i>	<i>R</i> ²	ΔR^2	<i>R</i>	<i>R</i> ²	ΔR^2
Statistically Significant Broad Index Score Variability (<i>n</i> = 456)						
GCA	.58	.34	.34*	.59	.34	.34*
+ Profile Group	.58	.34	.00	.60	.36	.02*
+ GCA × Group	.58	.34	.00	.60	.36	.00
15% Base Rate (<i>n</i> = 332)						
GCA	.57	.32	.32*	.60	.36	.36*
+ Profile Group	.57	.33	.01	.61	.37	.01
+ GCA × Group	.57	.33	.00	.61	.37	.00
10% Base Rate (<i>n</i> = 246)						
GCA	.54	.29	.29*	.42	.18	.18*
+ Profile Group	.54	.29	.00	.43	.19	.01
+ GCA × Group	.54	.30	.01	.44	.19	.00
5% Base Rate (<i>n</i> = 130)						
GCA	.53	.28	.28*	.61	.38	.38*
+ Profile Group	.54	.29	.01	.62	.38	.00
+ GCA × Group	.54	.29	.00	.62	.38	.00

Note. GCA = General Conceptual Ability.
 **p* < .001.

mathematically derived full-scale IQ, will more accurately reflect the person's learning abilities. When there is a marked discrepancy across verbal and performance scores, averaging to obtain a full-scale IQ score can be misleading. (p. 42)

Therefore, the findings of the current study have the potential to significantly affect several uses of IQ measures, including educational placements and psychoeducational diagnoses.

Despite these findings, limitations to this study exist. First, additional research is needed to determine if these findings replicate studies with other ability and achievement test batteries, especially with such recent revisions as the DAS-II and Stanford-Binet—Fifth Edition (Roid, 2003). Examining additional IQ batteries, such as the SB-V, is important in order to determine whether equality of *g* loadings and

specificity across broad index scores affects the predictability of the GCA and broad index scores. However, Watkins, Glutting, and Lei (2007) investigated the predictive validity of IQs in the presence of factor discrepancies on the WISC-III and the Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003). Similar results were found in that the WISC-III and Wechsler Intelligence Scale for Children—Fourth Edition IQs were unbiased predictors of academic achievement in the presence of an uneven factor score profile. This lends support to the application of the present results to other cognitive ability measures until further research on specific tests is conducted.

Another limitation of this study is found within the sample of participants. First, the sample sizes in the analyses of broad index

Table 9
Summary of Regression Results for Verbal Broad Index Score to Each Academic Area Across Prevalence Levels

Test	Verbal Broad Index to Basic Number Skills			Verbal Broad Index to Word Reading		
	<i>R</i>	<i>R</i> ²	ΔR^2	<i>R</i>	<i>R</i> ²	ΔR^2
Statistically Significant Broad Index Score Variability (<i>n</i> = 382)						
GCA	.52	.27	.27*	.52	.27	.27*
+ Profile Group	.52	.27	.00	.52	.27	.00
+ GCA \times Group	.52	.27	.00	.52	.27	.00
15% Base Rate (<i>n</i> = 276)						
GCA	.53	.28	.28*	.47	.22	.22*
+ Profile Group	.53	.28	.00	.48	.23	.01
+ GCA \times Group	.53	.28	.00	.49	.24	.01
10% Base Rate (<i>n</i> = 188)						
GCA	.54	.29	.29*	.42	.18	.18*
+ Profile Group	.54	.29	.00	.43	.19	.01
+ GCA \times Group	.54	.30	.01	.44	.19	.00
5% Base Rate (<i>n</i> = 114)						
GCA	.55	.30	.30*	.38	.14	.14*
+ Profile Group	.55	.30	.00	.39	.15	.01
+ GCA \times Group	.56	.31	.01	.41	.17	.02

Note. GCA = General Conceptual Ability.

**p* < .001.

discrepancies occurring rarely in the population were relatively small. A larger sample size at these rare levels might have resulted in significance. Therefore, further research with larger samples is needed to determine if the IQ is statistically biased in the presence of extreme broad index score differences. Second, the sample was collected in the early 1990s, and therefore future research should examine more recent samples. Finally, the sample was normative and did not allow for additional analyses for children with disabilities. Thus, future research should examine the generalizability of these results to children with learning disabilities, mental retardation, and so on.

In addition, although attempts were made to match the discrepant and nondiscrepant groups on demographic variables, imperfect results were obtained. The two groups were

matched successfully on GCA and demographic variables as demonstrated by nonsignificant statistical tests, but they were not perfectly matched on demographics. It is possible that these small differences in demographic characteristics contributed to the findings.

Finally, the DAS ability and achievement tests were conformed and might, therefore, share variance unrelated to the constructs being measured. If true, this should artificially inflate the correlations between DAS ability and achievement scores. However, it was found that the GCA of the matched groups was moderately correlated with achievement scores ($r = .52-.57$), which is similar to the IQ-achievement correlation of .60 recommended by Sattler (2001). Thus, there is little evidence of variance contamination, and results from nonrelated IQ and achievement

tests should produce the same results, albeit with reduced predictive accuracy.

Implications for Practice

Although the results of the current study indicate that the IQ is the most valid predictor of achievement, there are conditions under which the IQ is invalid. For example, if a portion of IQ subtests measure construct-irrelevant influences (e.g., hearing, vision, language, or attention problems) or were prone to error, the IQ would not be a valid indicator of academic performance. Therefore, the IQ can only be considered the most valid indicator of achievement if the administration is valid and if the examinee's responses reflect the intended construct. It is only under these standardized conditions to which the results of this study speak.

Psychometrically, the variance shared by subtests can be attributed to first-order factors or to the general factor (Gignac, 2006). A preference for the one or the many should be based on such scientific principles as replication and parsimony. In the current study, the IQ retained its predictive accuracy in the presence of significant interfactor variability. Other research has demonstrated similar results with other IQ batteries (Glutting et al., 1997; Kline et al., 1993; Watkins et al., 2007; Youngstrom et al., 1999), and general intelligence has shown robust predictive power across broad domains of human functioning (Gottfredson, 1997, 2004; Lubinski, 2000; Schmidt & Hunter, 2004). Given these results, the law of parsimony suggests that interpretation of intelligence tests should focus on the IQ because "it is vain to do with more what can be done with less" (Jeffreys & Berger, 1992, p. 64). Thus, the interpretation of profiles of broad index scores will not contribute any additional prediction of achievement. Although further research is needed, competent psychological practice should be based on the best available evidence (Gibbs, 2003). Clinicians should, therefore, be extremely reluctant to discount the predictive validity of the IQ even in the presence of uneven broad index score profiles.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. New York: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Arlington, VA: American Psychiatric Press.
- American Psychological Association. (2002). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.
- Aylward, G. P. (1992). Review of the Differential Ability Scales. In J. J. Kramer & J. C. Conoley (Eds.), *The eleventh mental measurements yearbook* (pp. 281–282). Lincoln: The Buros Institute of Mental Measurements of the University of Nebraska—Lincoln.
- Braden, J. P. (1997). The practical impact of intellectual assessment issues. *School Psychology Review*, 26, 242–248.
- Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- DiCerbo, K. E., & Barona, A. (2000). A convergent validity study of the Differential Ability Scales and the Wechsler Intelligence Scale for Children—Third Edition with Hispanic children. *Journal of Psychoeducational Assessment*, 18, 344–352.
- Donders, J. (1996). Cluster subtypes in the WISC-III standardization sample: Analysis of factor index scores. *Psychological Assessment*, 8, 312–318.
- Dori, G. A., & Gordon, C. J. (2004). Education-stratified base-rate information on discrepancy scores within and between the Wechsler Adult Intelligence Scale—Third Edition and the Wechsler Memory Scale—Third Edition. *Psychological Assessment*, 16, 146–154.
- Drummond, R. J. (2004). *Appraisal procedures for counselors and helping professionals* (5th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Elliott, C. D. (1990a). *Differential Ability Scales administration and scoring manual*. New York: Harcourt Brace Jovanovich.
- Elliott, C. D. (1990b). *Differential Ability Scales introductory and technical handbook*. New York: Harcourt Brace Jovanovich.
- Elliott, C. D. (2005). The Differential Ability Scales. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed.; pp. 402–424). New York: Guilford.
- Elliott, C. D. (2006). *Differential Ability Scales—Second Edition*. New York: Harcourt Brace Jovanovich.
- Gibbs, L. E. (2003). *Evidence-based practice for the helping professions: A practical guide with integrated multimedia*. Pacific Grove, CA: Brooks/Cole.
- Gignac, G. E. (2006). The WAIS-III as a nested factors model: A useful alternative to the more conventional oblique and higher-order models. *Journal of Individual Differences*, 27, 73–86.
- Glutting, J. J., McDermott, P. A., Konold, T. R., Snelbaker, A. J., & Watkins, M. W. (1998). More ups and downs of subtest analysis: Criterion validity of the DAS with an unselected cohort. *School Psychology Review*, 27, 599–612.

- Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAT-II. *The Journal of Special Education, 40*, 103–114.
- Glutting, J. J., Youngstrom, E. A., Ward, T., Ward, S., & Hale, R. L. (1997). Incremental efficacy of WISC-III factor scores in predicting achievement: What do they tell us? *Psychological Assessment, 9*, 295–301.
- Gottfredson, L. S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence, 24*, 79–132.
- Gottfredson, L. S. (2004). Schools and the *g* factor. *Wilson Quarterly, Summer*, 35–45.
- Hale, J. B., & Fiorello, C. A. (2001). Beyond the academic rhetoric of 'g': Intelligence testing guidelines for practitioners. *The School Psychologist, 55*, 113–135.
- Hale, J. B., Fiorello, C. A., Kavanagh, J. A., Hoepfner, J. B., & Gaither, R. A. (2001). WISC-III predictors of academic achievement for children with learning disabilities: Are global and factor scores compatible? *School Psychology Quarterly, 16*, 31–55.
- Individuals with Disabilities in Education Improvement Act. (2004). Public Law 108–446 (20 U.S.C. 1400 *et seq.*). Retrieved September 14, 2007, from <http://idea.ed.gov/download/finalregulations.html>.
- Jeffreys, W. H., & Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *American Scientist, 80*, 64–72.
- Kamphaus, R. W. (2001). *Clinical assessment of child and adolescent intelligence* (2nd ed.). Boston: Allyn & Bacon.
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.
- Kaufman, A. S., & Lichtenberger, E. O. (2000). *Essentials of WISC-III and WPPSI-R assessment*. New York: Wiley.
- Kaufman, A. S., & Lichtenberger, E. O. (2002). *Assessing adolescent and adult intelligence* (2nd ed.). Boston: Allyn & Bacon.
- Kline, R. B., Snyder, J., Guilmette, G., & Castellanos, M. (1993). External validity of the profile variability index for the K-ABC, Stanford-Binet, and WISC-R: Another cul-de-sac. *Journal of Learning Disabilities, 26*, 557–567.
- Kraemer, H. C., & Blasey, C. M. (2004). Centering in regression analyses: A strategy to prevent errors in statistical inference. *International Journal of Methods in Psychiatric Research, 13*, 141–151.
- Kubiszyn, T. W., Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., & Eisman, E. J. (2000). Empirical support for psychological assessment in clinical health care settings. *Professional Psychology: Research and Practice, 31*, 119–130.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.
- Lichtenberger, E. D., Kaufman, A. S., & Lai, Z. C. (2002). *Essentials of WMS-III assessment*. New York: Wiley.
- Lubinski, D. (2000). Scientific and social significance of assessing individual differences: "Sinking shafts at a few critical points." *Annual Review of Psychology, 51*, 405–444.
- McDermott, P. A., & Glutting, J. J. (1997). Informing stylistic learning behavior, disposition, and achievement through ability subtests: Or more illusions of meaning? *School Psychology Review, 26*, 163–175.
- McDermott, P. A., Marston, N. C., & Stott, D. H. (1993). *Adjustment Scales for Children and Adolescents*. Philadelphia: Edumatic and Clinical Science.
- McIntosh, D. E. (1999). Identifying at-risk preschoolers: The discriminant validity of the upper preschool level of the Differential Ability Scales (DAS). *Psychology in the Schools, 36*, 1–10.
- Oh, H. J., Glutting, J. J., Watkins, M. W., Youngstrom, E. A., & McDermott, P. A. (2004). Correct interpretation of latent versus observed abilities: Implications from structural equation modeling applied to the WISC-III and WIAT linking sample. *Journal of Special Education, 38*, 159–173.
- Pfeiffer, S. I., Reddy, L. A., Kletzel, J. E., Schmelzer, E. R., & Boyer, L. M. (2000). The practitioner's view of IQ testing and profile analysis. *School Psychology Quarterly, 15*, 376–385.
- Prifitera, A., Saklofske, D. H., & Weiss, L. G. (2005). *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives*. New York: Elsevier Academic Press.
- Reinehr, R. C. (1992). Review of the Differential Ability Scales. In J. J. Kramer & J. C. Conoley (Eds.), *The eleventh mental measurements yearbook* (pp. 282–283). Lincoln: The Buros Institute of Mental Measurements of the University of Nebraska—Lincoln.
- Reynolds, C. R. (1995). Test bias and the assessment of intelligence and personality. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 545–573). New York: Plenum.
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales—Fifth Edition*. Itasca, IL: Riverside Publishing.
- Ryan, J. J., Kreiner, D. S., & Burton, D. B. (2002). Does high scatter affect the predictive validity of WAIS-III IQs? *Applied Neuropsychology, 9*, 173–178.
- Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Author.
- Sattler, J. M., & Dumont, R. (2004). *Assessment of children: WISC-IV and WPPSI-III Supplement*. San Diego, CA: Jerome M. Sattler Publisher.
- Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*, 162–173.
- Society for Industrial and Organization Psychology. (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author.
- Stone-Romero, E. F., & Anderson, L. E. (1994). Relative power of moderated multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects. *Journal of Applied Psychology, 79*, 354–359.
- Thorndike, R. L., Hagen, E., & Sattler, J. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago: Riverside Publishing Company.
- U.S. Department of Commerce. (1990). *School enrollment—Social and economic characteristics of students* (Current Population Reports, Series P-20, No. 443). Washington, DC: U.S. Bureau of the Census.
- Vaughn, S., Gersten, R., & Chard, D. J. (2000). The underlying message in LD intervention research: Findings from research syntheses. *Exceptional Children, 67*, 99–114.

- Watkins, M. W., & Glutting, J. J. (2000). Incremental validity of WISC-III profile elevation, scatter, and shape information for predicting reading and math achievement. *Psychological Assessment, 12*, 402–408.
- Watkins, M. W., Glutting, J. J., & Lei, P. W. (2007). Validity of the Full Scale IQ when there is significant variability among WISC-III and WISC-IV factor scores. *Applied Neuropsychology, 14*, 13–20.
- Watkins, M. W., Glutting, J. J., & Youngstrom, E. A. (2005). Issues in subtest profile analysis. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 251–268). New York: Guilford.
- Watkins, M. W., & Kush, J. C. (1994). Wechsler subtest analysis: The right way, the wrong way, or no way? *School Psychology Review, 23*, 640–651.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children—Revised*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale—Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale—Third Edition*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2001). *Wechsler Individual Achievement Test—Second Edition*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—Fourth Edition*. San Antonio, TX: Psychological Corporation.
- Weiss, L. G., Saklofske, D. H., & Prifitera, A. (2005). Interpreting the WISC-IV index scores. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 71–100). New York: Elsevier Academic Press.
- Youngstrom, E. A., Kogos, J. L., & Glutting, J. J. (1999). Incremental efficacy of Differential Ability Scales factor scores in predicting individual achievement criteria. *School Psychology Quarterly, 14*, 26–39.

Date Received: February 2, 2007

Date Accepted: November 11, 2007

Action Editor: Shane Jimerson ■

Kasey M. Kotz received her MS in school psychology from the Pennsylvania State University, and she was subsequently accepted as a doctoral candidate in the School Psychology Program at the same institution. She is currently completing her predoctoral internship at Pennridge School District in eastern Pennsylvania.

Marley W. Watkins, PhD, is training director and professor in the School Psychology Program at Arizona State University and a diplomate of the American Board of Professional Psychology. He received his PhD in school psychology from the University of Nebraska—Lincoln and was a practicing school psychologist for 15 years before entering academe.

Paul A. McDermott, PhD, is a professor with the Graduate School of Education and the School of Medicine at the University of Pennsylvania. He is a school psychologist specializing in measurement and multivariate statistics. His primary affiliation is with the PhD Program in Policy Research, Evaluation, and Measurement, where he teaches item response theory and multilevel modeling, codirects randomized controlled trials for Head Start, and studies neighborhood effects on school outcomes.