Applied Neuropsychology
2007, Vol. 14, No. 1, 13–20

Validity of the Full-Scale IQ When There Is Significant Variability Among WISC-III and WISC-IV Factor Scores

Marley W. Watkins
Pennsylvania State University, University Park, Pennsylvania, USA
Joseph J. Glutting
University of Delaware, Newark, Delaware, USA
Pui-Wa Lei
Pennsylvania State University, University Park, Pennsylvania, USA

Fiorello et al. (this issue) asserted that as subtest or factor-score variability increases, the overall ability score becomes a less viable predictor of achievement—and that this is especially the case for clinical populations. The present study tested the contention that global IQ scores are invalid predictors of achievement in the presence of significant variability among constituent factors. It did so using three samples of participants: (a) 412 from the WISC-III/WIAT linking sample, (b) 460 enrolled in special education programs, and (c) 136 from the WISC-IV/WIAT-II linking sample. In each sample, participants with and without statistically significant factor score variability were matched on Full Scale IQ (FSIQ), age, gender, and race/ethnicity. The special education sample was also matched on disability category. For all samples, the FSIQ was a significant predictor of performance on reading and mathematics tests, but neither factor score variability nor the interaction of the FSIQ and factor score variability made a statistically significant incremental contribution to the prediction of reading and mathematics scores. Thus, the FSIQ was a robust predictor of achievement in regular and clinical samples, regardless of factor variability. Results are discussed in terms of the diagnostic importance of factor- and subtest-score variability.

Extensive research on IQ tests has verified the utility of general intelligence as a predictor of academic achievement (Kubiszyn et al., 2000; Lubinski, 2000). For example, the median correlation between global intelligence and contemporary achievement tests is .64 (Naglieri & Bornstein, 2003). Given this ubiquitous relationship, IQ tests are included in millions of child assessments each year (Kamphaus, Petoskey, & Rowe, 2000; Sattler, 2001), where global IQs are used as "an estimate of current and likely future academic performance" (Reschly & Grimes, 1990, p. 434).

However, many psychologists believe that while the global IQs are "useful summaries" (Kaufman & Lichtenberger, 2002, p. 47), they are rendered uninterpretable by significant variability among their subcomponents (Flanagan & Kaufman, 2004; Kaufman, 1994; Lezak, 1995; Weiss, Saklofski, & Prifitera, 2003; Wolber & Carne, 2002). More radically, some authors have declared global IQs to be unreliable and invalid for predicting academic achievement in the presence of intracognitive variability (Fiorello, Hale, McGrath, Ryan, & Quinn, 2002; Hale, Fiorello, Bertin, & Sherman, 2003), and have encouraged practitioners to "*never interpret the global IQ score if there is significant scatter or score variability*" (Hale & Fiorello, 2001, p. 132). This sweeping recommendation was based on regression communality analysis of global and factor index scores from the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991), and achievement scores from the Wechsler Individual Achievement Test (WIAT; Wechsler, 1992).

Fiorello and colleagues (Fiorello, Hale, Holdnack, Kavanagh, Terrell, & Long, this issue) contend that regression communality analysis identifies the proportion of variance in a dependent variable that

Address correspondence to Marley W. Watkins, Ph.D., Pennsylvania State University, 125 CEDAR Building, University Park, PA 16802. E-mail: mww10@psu.edu

can be uniquely attributed to each independent variable. Therefore, they are "determining the relative importance of independent variables" (Pedhazur, 1997, p. 243), an explanatory purpose for which regression communality analysis is ill suited (Cohen, Cohen, West, & Aiken, 2003; Lindenberger & Potter, 1998). Pedhazur (1997) declared that "the incremental partitioning of variance is *not* a valid approach for determining the relative importance of variables" (p. 245) and suggested that "it is informative only for predictive research, where it can be used for decisions about which variables may be eliminated while sacrificing little in overall predictability" (p. 269). He wryly noted that "when applying commonality analysis...for explanatory purposes, various authors refer the reader to the second edition of this book, without the slightest hint that I argued (strongly, I believe) against its use for this very purpose" (Pedhazur, 1997, p. 243).

Explanatory and predictive research "involve different research questions and study designs, different inferential approaches, different analysis strategies, and different reported information" (Huberty, 2003, p. 271). "Predictive research emphasizes practical applications, whereas explanatory research focuses on achieving a theoretical understanding of the phenomenon of interest" (Venter & Maxwell, 2000, p. 152). Following these guidelines, employing multiple regression to identify the most parsimonious predictor(s) of achievement is appropriate (see Glutting, Youngstrom, Ward, Ward, & Hale, 1997), but applying multiple regression to draw conclusions regarding the meaningfulness of IQ variables is inappropriate. The latter task, determining the relative proportion of common score variance among IQ subtests and factors, is best accomplished with factor analysis (Gustafsson & Undheim, 1996; McDonald, 1985), not regression analysis.

It might be argued that the studies conducted by Fiorello et al. (this issue, 2002) were predictive in nature and therefore communality analysis was properly applied. However, their communality analyses were not used to select the most parsimonious predictors among the alternatives in question (FSIQ and factor scores). Rather, they treated FSIQ as the criterion measure and partitioned variance among its constituent predictor factors to identify the "structure of cognitive functioning" (Fiorello et al., this issue), an explanatory purpose for which communality analysis is unsuited (Pedhazur, 1997).

Regardless, practitioners are often advised that significant variability among subtest or factor scores renders the global IQ score an inaccurate estimate of overall ability (Flanagan & Kaufman, 2004; Kaufman, 1994; Lezak, 1995; Weiss, Saklofski, & Prifitera, 2003; Wolber & Carne, 2002), and, consequently, an invalid predictor of academic achievement (Gridley & Roid, 1998; Hildebrand & Ledbetter, 2001). In those cases, psychologists are encouraged to disregard the global IQ and use factor or subtest scores to predict academic achievement. However, this hypothesis has not been directly tested. Is the global IQ a valid predictor of academic achievement among flat cognitive profiles and an invalid predictor of academic achievement when constituent scores are variable? The present study directly addressed that question.

METHOD

Instruments

The WISC-III (Wechsler, 1991) is an individually administered test of intelligence for children aged 6 years through 16 years, 11 months that was standardized on a nationally representative sample ($N = 2{,}200$) closely approximating the 1988 United States Census on gender, parent education, race/ethnicity, and geographic region. The WISC-III has 13 individual subtests ($M = 10$, $SD = 3$), ten standard and three supplementary, that combine to yield three composites: Verbal (VIQ), Performance (PIQ), and Full Scale (FSIQ) IQs ($M = 100$, $SD = 15$). In addition, the WISC-III provides four factor-based indexes: Verbal Comprehension (VC), Perceptual Organization (PO), Freedom from Distractibility (FD), and Processing Speed (PS) ($M = 100$, $SD = 15$). Full details of the WISC-III and its standardization are presented in Wechsler (1991). Additional reliability and validity data are provided by Sattler (2001) as well as Zimmerman and Woo-Sam (1997).

The Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) is a revised and updated version of the WISC-III. It was standardized on a nationally representative sample ($N = 2{,}200$) of children aged 6–16 years closely approximating the 2000 United States Census on gender, race, parent education level, and geographic region. The WISC-IV has 10 core subtests ($M = 10$,

$SD = 3$) that form four factor indices ($M = 100$, $SD = 15$): Verbal Comprehension (VC), Perceptual Reasoning (PR), Working Memory (WM), and Processing Speed (PS). The FSIQ is based on the sum of scores from the 10 core subtests.

The WIAT (Wechsler, 1992) is an individually administered test of academic achievement which was standardized with 4,252 children in grades K–12. This test contains nine subtests that are aggregated into four composite scores: reading, mathematics, language, and writing ($M = 100$; $SD = 15$). For this study, the reading and mathematics composite scores were used.

The Wechsler Individual Achievement Test-Second Edition (WIAT-II, Wechsler, 2002) is a revised and expanded version of the WIAT. It was standardized on 5,586 individuals in grades pre-kindergarten through 16 closely approximating the 1998 United States Census on grade, age, gender, race/ethnicity, geographic region, and parent education level. The WIAT-II contains nine subtests that combine to form four composite scores: reading, mathematics, written language, and oral language ($M = 100$; $SD = 15$). For this study, the reading and mathematics composite scores were used. Extensive reliability and validity evidence for the WIAT and WIAT-II was provided by Hendry (2003), Smith (2001), and Wechsler (2002).

Participants/Procedures

Sample 1

The WISC-III/WIAT linking sample of 550 male and 568 female students (M age = 10.9, $SD = 3.1$) comprised the first group of participants. Ethnicity was 76% White, 12% Black, 10% Hispanic, and 2% other. As expected, their WISC-III scores were average (M FSIQ = 100, VIQ = 99, PIQ = 101). Complete details of this sample are provided in Wechsler (1992).

These 1,118 children and adolescents were divided into two groups based upon their degree of WISC-III factor score variability. There were 912 individuals with at least one statistically significant factor score difference and 206 youth without significant factor score variability (Wechsler, 1991, p. 261). Each of the 206 individuals with a flat factor score profile was then sequentially matched as closely as possible to a participant from the variable profile group on FSIQ, age, gender, and ethnicity. One case could not be matched on FSIQ and was discarded, leaving a final sample of 412 participants (210 male and 200 female). Ethnicity was 73% White, 14% Black, 11% Hispanic, and 2% other.

Sample 2

Requests to contribute to an investigation of the WISC-III were mailed to 9,227 school psychologists throughout the United States. Practitioners were invited to participate; all worked in a school setting and were members of the National Association of School Psychologists. Those school psychologists were asked to report anonymous data from their five most recent evaluations that resulted in special education placement under the following categories: learning disability (LD), emotional disability (ED), or mental retardation (MR). Full details of this sample are provided in Watkins (2005).

Anonymous scores were reported on 2,356 participants; however, 1,112 of those individuals could not be retained due to missing data on intelligence, achievement, or demographic variables. The 1,244 remaining participants included 842 males and 402 females (M age = 10.0, $SD = 2.6$). Ethnicity was 76% White, 13% Black, 7% Hispanic, and 4% other. Special education enrollment was 70% LD, 8% MR, 11% ED, and 11% in other or multiple categories. Their WISC-III scores were within the low average to average range (M FSIQ = 92, VIQ = 91, PIQ = 94).

These 1,244 participants were divided into two groups based upon their degree of WISC-III factor score variability. There were 1,014 individuals with at least one statistically significant factor score difference and 230 without significant factor score variability (Wechsler, 1991, p. 261). Each of the 230 participants with a flat factor score profile was then sequentially matched as closely as possible to a participant from the variable profile group on disability category (including type of LD, if applicable; reading, math, writing), FSIQ, age, gender, and ethnicity. Unfortunately, 24 of the youth with flat WISC-III factor score profiles could not be matched on disability category or FSIQ and were discarded, leaving a final sample of 412 (279 male and 133 female, M age = 10.3, $SD = 2.6$). Ethnicity was 76% White, 14% Black, 6% Hispanic, and 4% other. Special education enrollment was 66% LD, 15% MR, 16% ED, and 3% in other or multiple categories.

Academic achievement of the participants enrolled in special education was measured by a total of 62 tests or combinations of tests. However, versions of the Woodcock-Johnson achievement test and the WIAT were used in around 85% of the cases. Achievement in reading and math was first determined by averaging the reading achievement scores (i.e., basic reading skills and reading comprehension subtests) and math achievement scores (i.e., math computation and math reasoning subtests). If only one achievement score was provided in reading or math, that score was used.

Sample 3

The WISC-IV/WIAT-II linking sample of 282 males and 268 females (*M* age = 11.6, *SD* = 3.2) comprised the third group of participants. Ethnicity was 61% White, 16% Black, 18% Hispanic, and 3% other. As expected, their WISC-IV scores were average (*M* FSIQ = 100, VC = 100, PR = 99, WM = 99, PS = 101). Complete details of this sample are provided in Wechsler (2003).

These 550 youth were divided into two groups based upon their degree of WISC-IV factor score variability. There were 468 individuals with at least one statistically significant factor score difference and 82 without significant factor score variability (Wechsler, 2003, p. 256). Each of the 82 participants with a flat profile was then sequentially matched as closely as possible to a participant from the variable factor profile group on FSIQ, gender, ethnicity, and age. Unfortunately, 14 individuals were missing matching or criterion variable data, leaving a final sample of 136 participants (68 with flat factor index profiles and 68 with variable factor index profiles). This final sample contained 66 males and 70 females (*M* age = 11.2, *SD* = 3.1). Ethnicity was 70% White, 13% Black, 15% Hispanic, and 2% other.

Analyses

As recommended by the Society for Industrial and Organization Psychology (2003), moderated multiple regression was used to detect any bias in the predictive validity of FSIQ scores between participants with and without significant factor score variability (Stone-Romero & Anderson, 1994). Using this method, composite reading and mathematics scores were sequentially regressed on the FSIQ, factor profile group membership, and an interaction term between the FSIQ and profile group. A statistically significant main effect for profile group membership or interaction between FSIQ and profile group membership would signal differential predictive validity. Because multiple significance tests were conducted with overlapping participants, a conservative alpha level ($p < .01$) was adopted for each test to control the overall Type I error rate.

RESULTS

Descriptive statistics for intelligence and achievement scores for the three samples are provided in Tables 1–3, respectively. As expected, there were highly significant differences in the amount of factor variability between flat and variable profile groups in all three samples. Matching on FSIQs was successful as there were no statistically significant differences between flat and variable profile groups on the FSIQ. Nor was there any statistically significant differences in proportional membership in age, gender, race/ethnicity, or disability categories across flat and variable profile groups.

The correlation coefficients between the FSIQ and achievement scores for the flat and variable profile groups are presented in Table 4. None of the between-group coefficients were statistically significantly different ($p > .26$). Table 5 displays the sequential moderated multiple regression results. For all samples, FSIQs were significant predictors of performance on reading and mathematics tests, but neither factor profile group (flat versus variable) nor the interaction between the

Table 1. *Descriptive Statistics for WISC-III and WIAT Scores Among 205 Students with Flat Factor Profiles and 205 Students with Variable Factor Profiles from the WISC-III/WIAT Linking Sample*

	Flat Profiles		Variable Profiles	
Score	Mean	SD	Mean	SD
FSIQ	99.2	14.9	99.2	14.9
VC	99.7	13.2	97.7	17.4
PO	99.7	13.0	100.5	16.8
FD	99.1	13.0	100.6	17.3
PS*	99.5	13.0	105.5	16.7
Reading	99.9	14.5	100.6	15.1
Mathematics	100.4	15.1	101.1	15.7
Scatter*	9.3	2.9	33.0	9.1

Note. FSIQ = Full Scale IQ, VC = Verbal Comprehension factor score, PO = Perceptual Organization factor score, FD = Freedom from Distractibility factor score, PS = Perceptual Speed factor score, Scatter = absolute value of largest factor score difference.
*$p < .001$.

Table 2. *Descriptive Statistics for WISC-III and WIAT Scores Among 206 Students with Flat Factor Profiles and 206 Students with Variable Factor Profiles from the Special Education Sample*

	Flat Profiles		Variable Profiles	
Score	**Mean**	**SD**	**Mean**	**SD**
FSIQ	87.8	16.3	87.8	15.8
VC	90.0	14.4	89.3	15.5
PO	90.3	15.1	91.9	17.9
FD	88.4	13.7	85.4	13.5
PS	90.1	14.4	91.2	17.3
Reading	85.0	14.7	85.6	13.6
Mathematics	87.3	15.1	87.0	15.0
Scatter*	9.9	3.1	25.4	8.1

Note. FSIQ = Full Scale IQ, VC = Verbal Comprehension factor score, PO = Perceptual Organization factor score, FD = Freedom from Distractibility factor score, PS = Perceptual Speed factor score, and Scatter = absolute value of largest factor score difference.
*$p < .001$.

Table 4. *Correlation Coefficients Between Full Scale IQ and Achievement Composite Scores (Reading and Mathematics) for Students With and Without Significant Factor Score Variability in the WISC-III/WIAT Linking Sample, Special Education Sample, and WISC-IV/WIAT-II Linking Sample*

	Reading		Math	
Sample	**Flat Profile**	**Variable Profile**	**Flat Profile**	**Variable Profile**
WISC-III/WIAT Link Sample[a]	.72	.65	.78	.61
Special Education Sample[b]	.61	.54	.80	.74
WISC-IV/WIAT-II Link Sample[c]	.76	.77	.81	.67

[a]205 students in each profile group.
[b]206 students in each profile group.
[c]68 students in each profile group.

FSIQ and factor profile group significantly added to the prediction. Thus, there were no statistically significant differences between participants with and without significant factor variability once FSIQ was controlled. Importantly, the factor profile group main effects and its interaction with FSIQ were *not* practically significant: no R^2 increment exceeded .02 and most were less than .01.

DISCUSSION

Some psychologists have claimed that the global IQ is invalid in the presence of significant variability among constituent subtest and/or factor scores (Fiorello et al., 2002, this issue; Hale & Fiorello, 2001; Hale et al., 2003). This contention was tested in the present study by matching children and adolescents on global IQ (i.e., FSIQ) and then testing whether the FSIQ predicted performance on reading and mathematics tests equally well for participants with and without

Table 3. *Descriptive Statistics for WISC-IV and WIAT-II Scores Among 68 Students with Flat Factor Profiles and 68 Students with Variable Factor Profiles from the WISC-IV/WIAT-II Linking Sample*

	Flat Profiles		Variable Profiles	
Score	**Mean**	**SD**	**Mean**	**SD**
FSIQ	101.7	14.3	101.7	14.3
VC	101.1	11.8	101.0	14.7
PR	100.5	11.3	101.6	14.3
WM	101.0	11.5	99.0	15.2
PS	101.0	11.7	102.5	14.6
Reading	102.7	15.7	98.4	15.4
Mathematics	103.4	14.9	102.8	16.9
Scatter*	8.3	2.6	23.8	7.9

Note. FSIQ = Full Scale IQ, VC = Verbal Comprehension factor score, PR = Perceptual Reasoning factor score, WM = Working Memory factor score, PS = Perceptual Speed factor score, and Scatter = absolute value of largest factor score difference.
*$p < .001$.

Table 5. *Regression of Reading and Mathematics Scores on FSIQ, Profile Group Membership, and the Interaction of Full Scale IQ and Profile Group Membership for Students in the WISC-III/WIAT Linking Sample, Special Education Sample, and WISC-IV/WIAT-II Linking Sample*

	Reading		Mathematics	
Sample/Entry Order	**R^2**	**R^2 Change**	**R^2**	**R^2 Change**
WISC-III/WIAT link sample[a]				
FSIQ	.472	.472*	.482	.482*
+Profile group	.473	.001	.483	.001
+FSIQ × Profile group	.473	.000	.488	.005
Special education sample[b]				
FSIQ	.335	.335*	.591	.591*
+Profile group	.336	.001	.591	.000
+FSIQ × Profile group	.338	.002	.592	.001
WISC-IV/WIAT-II link sample[c]				
FSIQ	.572	.572*	.644	.644*
+Profile group	.592	.020	.644	.000
+FSIQ × Profile group	.592	.000	.654	.009

Note. FSIQ = Full Scale IQ.
[a]205 students in each profile group.
[b]206 students in each profile group.
[c]68 students in each profile group.
*$p < .001$.

significant factor score variability. Results were unambiguous: the FSIQ was a robust predictor of academic achievement regardless of significant factor score variability for participants with and without disabilities. Thus, the hypothesis that global ability is invalid in the presence of factor variability was refuted.

Fiorello et al. (this issue) were originally invited to address the issue of subtest interpretation of Wechsler scales, a practice they advocate in their book (Hale & Fiorello, 2004). Unfortunately, their present study offered no empirical data on the topic, and therefore, we were compelled to address the issue of factor variability. With respect to their original charge, Fiorello et al. merely stated that "there is a dearth of empirically valid evidence" to support inferences that subtest scores may not be valid for clinical populations (this issue). To the contrary, empirical studies across the past 20 years have repeatedly demonstrated that subtest scores retain limited external validity. Examples of diminished utility include the inability of either individual subtest scores or score patterns to inform the identification of neurological deficits (Watkins, 1996), the diagnosis of learning disabilities (Daley & Nagle, 1996; Glutting, McGrath, Kamphaus, & McDermott, 1992; Kline et al., 1992; Maller & McDermott, 1997; Mueller, Dennis, & Short, 1986; Smith & Watkins, 2004; Ward, Ward, Hatt, Young, & Mollner, 1995; Watkins, 1999, 2000, 2003, 2005; Watkins & Kush, 1994; Watkins, Kush, & Glutting, 1997a, 1997b; Watkins, Kush, & Schaefer, 2002; Watkins & Worrell, 2000), or the classification of social, behavioral, and/or emotional problems (Beebe, Pfiffner, & McBurnett, 2000; Dumont, Farr, Willis, Whelley, 1998; Glutting et al., 1992, Glutting, McDermott, Konold, Snelbaker, & Watkins, 1998; Lipsitz, Dworkin, & Erlenmeyer-Kimling, 1993; McDermott & Glutting, 1997; Reinecke, Beebe, & Stein, 1999; Riccio, Cohen, Hall, & Ross, 1997; Rispens et al., 1997; Teeter & Korducki, 1998).

The claim that significant subtest score variability compromises the predictive validity of the global IQ was also tested by Ryan, Kreiner, and Burton (2002) with 80 adult medical patients with and without extreme intersubtest variability on the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997a). Participants were separated into high ($n = 40$) and low ($n = 40$) subtest variability groups and then matched on WAIS-III FSIQs. Wechsler Memory Scale-Third Edition (WMS-III; Wechsler, 1997b) scores served as criteria. Results showed that estimation of WMS-III scores was equally accurate for both high and low subtest variability groups. Based on these results, Ryan et al. concluded that "marked intersubtest scatter [does not] reduce the predictive validity of the FSIQ" (p. 177).

Our results challenge the practice of discounting the global IQ as a predictor of academic achievement when factor scores significantly vary. As anticipated by Kehle, Clark, and Jenson (1993), the "predictive validity afforded by the single full-scale score [was not] appreciably altered" (p. 151) by factor variability. These results are also consonant with the conclusions of Gustafsson and Undheim (1996) who found that due to aggregation, the general intelligence factor accounted for a substantial proportion of FSIQ variance (i.e., 71% for the WISC-III). However, it is possible that the global IQ would become less robust if there was more extreme variability among constituent factors than employed in the present study. This is an empirical question that future research should address (Sattler & Dumont, 2004), although evidence from extant investigations on the importance of cognitive scatter or variability is not promising (Glutting et al., 1997; Kline, Snyder, Guilmette, & Castellanos, 1992; Oh, Glutting, Watkins, Youngstrom, & McDermott, 2004; Watkins, 2005; Watkins & Glutting, 2000).

ACKNOWLEDGMENT

Note: Standardization data of the Wechsler Intelligence Scale For Children, Fourth Edition, Copyright © 2004, and the Wechsler Individual Achievement Test, Second Edition, Copyright © 2001, by Harcourt Assessment, Inc. Used with permission of the Publisher. All rights reserved. "Wechsler Intelligence Scale For Children," "WISC," "Wechsler Individual Achievement Test," and "WIAT" are trademarks of Harcourt Assessment, Inc., registered in the United States of America and/or other jurisdictions.

The authors thank Ms. Kasey Kotz for preparing the data files for the analyses reported in this paper. The authors are also grateful for a research award from the Society for the Study of School Psychology.

REFERENCES

Beebe, D. W., Pfiffner, L. J., & McBurnett, K. (2000). Evaluation of the validity of the Wechsler Intelligence Scale for Children, 3rd ed. Comprehension and picture arrangement subtests as measures of social intelligence. *Psychological Assessment, 12*, 97–101.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Daley, C. E., & Nagle, R. J. (1996). Relevance of WISC-III indicators for assessment of learning disabilities. *Journal of Psychoeducational Assessment, 14*, 320–333.

Dumont, R., Farr, L. P., Willis, J. O., & Whelley, P. (1998). 30-second interval performance on the Coding subtest of the WISC-III: Further evidence of WISC folklore? *Psychology in the Schools, 35*, 111–117.

Fiorello, C. A., Hale, J. B., Holdnack, J. A., Kavanagh, J. A., Terrell, J., & Long, L. (2007). Interpreting intelligence test results for children with disabilities: Is global intelligence relevant? *Applied Neuropsychology, 14*, 2–12.

Fiorello, C. A., Hale, J. B., McGrath, M., Ryan, K., & Quinn, S. (2002). IQ interpretation for children with flat and variable test profiles. *Learning and Individual Differences, 13*, 115–125.

Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. Hoboken, NJ: Wiley.

Glutting, J. J., McGrath, E. A., Kamphaus, R. W., & McDermott, P. A. (1992). Taxonomy and validity of subtest profiles on the Kaufman Assessment Battery for Children. *Journal of Special Education, 26*, 85–115.

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.

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.

Gridley, B. E., & Roid, G. H. (1998). The use of the WISC-III with achievement tests. In A. Prifitera & D. Saklofske (Eds.), *WISC-III clinical use and interpretation* (pp. 249–288). New York: Academic Press.

Gustafsson, J.-E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186–242). New York: MacMillan.

Hale, J. B., & Fiorello, C. A. (2001). Beyond the academic rhetoric of 'g': Intelligence testing guidelines for practitioners. *The School Psychologist, 55*(4), 113–139.

Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. New York: Guilford.

Hale, J. B., Fiorello, C. A., Bertin, M., & Sherman, R. (2003). Predicting math achievement through neuropsychological interpretation of WISC-III variance components. *Journal of Psychoeducational Assessment, 21*, 358–380.

Hendry, C. N. (2003). Clinical assessment of children's academic achievement. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence, aptitude, and achievement* (2nd ed.) (pp. 438–472). New York: Guilford.

Hildebrand, D. K., & Ledbetter, M. F. (2001). Assessing children's intelligence and memory: The wechsler intelligence scale for children-third edition and the children's memory scale. In J. J. W. Andrews, D. H. Saklofske, & H. L. Janzen (Eds.), *Handbook of psychoeducational assessment: Ability, achievement, and behavior in children* (pp. 13–32). New York: Academic Press.

Huberty, C. J. (2003). Multiple correlation versus multiple regression. *Educational and Psychological Measurement, 63*, 271–278.

Kamphaus, R. W., Petoskey, M. D., & Rowe, E. W. (2000). Current trends in psychological testing of children. *Professional Psychology: Research and Practice, 31*, 155–164.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.

Kaufman, A. S., & Lichtenberger, E. O. (2002). *Assessing adolescent and adult intelligence* (2nd ed.). Boston: Allyn and Bacon.

Kehle, T. J., Clark, E., & Jenson, W. R. (1993). The development of testing as applied to school psychology. *Journal of School Psychology, 31*, 143–161.

Kline, R. B., Snyder, J., Guilmette, S., & Castellanos, M. (1992). Relative usefulness of elevation, variability, and shape information from WISC-R, K-ABC, and fourth edition Stanford-Binet profiles in predicting achievement. *Psychological Assessment, 4*, 426–432.

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.

Lindenberger, U., & Potter, U. (1998). The complex nature of unique and shared effects in hierarchical linear regression: Implications for developmental psychology. *Psychological Methods, 3*, 218–230.

Lipsitz, J. D., Dworkin, R. H., & Erlenmeyer-Kimling, L. (1993). Wechsler comprehension and picture arrangement subtest and social adjustment. *Psychological Assessment, 5*, 430–437.

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.

Maller, S. J., & McDermott, P. A. (1997). WAIS-R profile analysis for college students with learning disabilities. *School Psychology Review, 26*, 575–585.

McDermott, P. A., & Glutting, J. J. (1997). Informing stylistic learning behavior, disposition, and achievement through ability subtest patterns: Practical implications for test interpretation. Special Issue. *School Psychology Review, 26*, 163–175.

McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Erlbaum.

Mueller, H. H., Dennis, S. S., & Short, R. H. (1986). A meta-exploration of WISC-R factor score profiles as a function of diagnosis and intellectual level. *Canadian Journal of School Psychology, 2*, 21–43.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment, 21*, 244–260.

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.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). New York: Harcourt Brace.

Reinecke, M. A., Beebe, D. W., & Stein, M. A. (1999). The third factor of the WISC-III: It's (probably) not freedom from distractibility. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 322–328.

Reschly, D. J., & Grimes, J. P. (1990). Best practices in intellectual assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-II* (pp. 425–439). Washington, DC: National Association of School Psychologists.

Riccio, C. A., Cohen, M. J., Hall, J., & Ross, C. M. (1997). The third and fourth factors of the WISC-III: What they don't measure. *Journal of Psychoeducational Assessment, 15*, 27–39.

Rispens, J., Swaab, H., van den Oord, E. J., Cohen-Kettenis, P., van Engeland, H., & van Yperen, T. (1997). WISC profiles in child psychiatric diagnosis: Sense or nonsense? *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1587–1594.

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: Jerome M. Sattler Publisher.

Sattler, J. M., & Dumont, R. (2004). *Assessment of children: WISC-IV and WPPSI-III supplement*. San Diego: Jerome M. Sattler.

Smith, D. R. (2001). Wechsler individual achievement test. In J. J. W. Andrews, D. H. Saklofske, & H. L. Janzen (Eds.), *Handbook of psychoeducational assessment: Ability, achievement, and behavior in children* (pp. 169–193). New York: Academic Press.

Smith, C. B., & Watkins, M. W. (2004). Diagnostic utility of the Bannatyne WISC-III pattern. *Learning Disabilities Research & Practice, 19*, 49–56.

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 moderate multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects. *Journal of Applied Psychology, 3*, 354–359.

Teeter, P. A., & Korducki, R. (1998). Assessment of emotionally disturbed children with the WISC-III. In A. Prifitera & D. H. Saklofske (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 119–138). New York: Academic Press.

Venter, A., & Maxwell, S. E. (2000). Issues in the use and application of multiple regression analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied statistics and mathematical modeling* (pp. 151–182). New York: Academic Press.

Ward, S. B., Ward, T. J., Hatt, C. V., Young, D. L., & Mollner, N. R. (1995). The incidence and utility of the ACID, ACIDS, and SCAD profiles in a referred population. *Psychology in the Schools, 32*, 267–276.

Watkins, M. W. (1999). Diagnostic utility of WISC-III subtest variability among students with learning disabilities. *Canadian Journal of School Psychology, 15*, 11–20.

Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly, 15*, 465–479.

Watkins, M. W. (2003). IQ subtest analysis: Clinical acumen or clinical illusion. *Scientific Review of Mental Health Practice, 2*, 118–141.

Watkins, M. W. (in press). Diagnostic validity of Wechsler subtest scatter. *Learning Disabilities: A Contemporary Journal.*

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., & Kush, J. C. (1994). WISC-R subtest analysis of variance: The right way, the wrong way, or no way? *School Psychology Review, 23*, 640–651.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997a). Discriminant and predictive validity of the WISC-III ACID profile among children with learning disabilities. *Psychology in the Schools, 34*, 309–319.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997b). Prevalence and diagnostic utility of the WISC-III SCAD profile among children with disabilities. *School Psychology Quarterly, 12*, 235–248.

Watkins, M. W., Kush, J. C., & Schaefer, B. A. (2002). Diagnostic utility of the learning disability index. *Journal of Learning Disabilities, 35*, 98–103.

Watkins, M. W., & Worrell, F. C. (2000). Diagnostic utility of the number of WISC-III subtests deviating from mean performance among students with learning disabilities. *Psychology in the Schools, 37*, 29–41.

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

Wechsler, D. (1992). *Wechsler individual achievement test*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997a). *Wechsler adult intelligence scale-third edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997b). *Wechsler memory scale-third edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2002). *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., Saklofski, D. H., & Prifitera, A. (2003). Clinical interpretation of the Wechsler intelligence scale for children-third edition (WISC-III) index scores. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence, aptitude, and achievement* (2nd ed.) (pp. 115–146). New York: Guilford.

Wolber, G. J., & Carne, W. F. (2002). *Writing psychological reports: A guide for clinicians* (2nd ed.). Sarasota, FL: Professional Resources Press.

Zimmerman, I. L., & Woo-Sam, J. M. (1997). Review of the criterion-related validity of the WISC-III: The first five years. *Perceptual and Motor Skills, 85*, 531–546.