Applied Neuropsychology
2007, Vol. 14, No. 1, 2–12

Interpreting Intelligence Test Results for Children with Disabilities: Is Global Intelligence Relevant?

Catherine A. Fiorello
Temple University, Philadelphia, Pennsylvania, USA
James B. Hale
Philadelphia College of Osteopathic Medicine, Philadelphia, Pennsylvania, USA
James A. Holdnack
The Psychological Corporation, San Antonio, Texas, USA
Jack A. Kavanagh
Loyola University, Chicago, Illinois, USA
Joy Terrell
Temple University, Philadelphia, Pennsylvania, USA
Lisa Long
Plattsburgh State University of New York, Plattsburgh, New York, USA

School psychological and neuropsychological evaluations typically include intellectual and other standardized assessment tools in the identification of children with disabilities. The clinical utility of intellectual assessment in the identification and treatment of these children has been repeatedly challenged, with alternatives such as a response to intervention or global intelligence score interpretation offered to replace the long-held tradition of idiographic interpretation of intellectual factors or subtests for the purpose of differential diagnosis and individualized intervention. Replicating previous work, this study examined the structure of intellectual functioning for children diagnosed with Learning Disability (LD; n = 128), Attention-Deficit/Hyperactivity Disorder (ADHD; n = 71), and traumatic brain injury (TBI; n = 29) using regression commonality analysis. Across groups, results provide substantial evidence for a multifactorial representation of intellectual functioning for children with LD, ADHD, or TBI, with little shared variance among factor predictors of FSIQ in each analysis. As global intellectual functioning, represented by the shared variance among all predictors, was largely absent and instead composed of several discrete elements with the requisite specificity for individual interpretation, idiographic interpretation appears to be warranted for children with disabilities.

Key words: ADHD, general intelligence, IQ, LD, profile analysis, TBI, test interpretation

In the early twentieth century, Alfred Binet developed a test to discriminate between children to determine which children would most benefit from individualized intervention (Neisser et al., 1996). Tests once used to determine individual capacity to achieve or function in academic institutions (Sternberg, Grigorenko, & Bundy, 2001) have been re-designed to identify children with learning disabilities (LD) and Attention-Deficit/Hyperactivity Disorder (ADHD), and, when combined with neuropsychological tests, traumatic brain injury (TBI). The use of standardized intellectual and cognitive tests for identification purposes has led to often contentious debate among researchers

---

Note: WISC-IV data were obtained with permission of Harcourt Assessment, Inc.

Address correspondence to Catherine A. Fiorello, Ph.D., Temple University (004-00), School Psychology Program, RA 269, 1301 Cecil B. Moore Ave., Philadelphia, PA 19122-6091. E-mail: catherine.fiorello@temple.edu

and practitioners alike regarding their utility in clinical practice.

This debate largely centers on the nature and structure of intellectual abilities. The early works of Horn and Cattell (1967) and Thurstone (1938) focused on multidimensional intellectual abilities, while Spearman (1904) emphasized a single general intellectual capacity, as suggested by a positive manifold among intellectual subtests. For children without disabilities, the positive manifold among intelligence subtests ensures single factor solutions or psychometric *g* can be derived from factor analyses (e.g., Jensen, 1998). Thought to be representative of *g*, global scores such as the Intelligence Quotient (IQ) have been used to predict meaningful life outcomes, such as educational attainment and occupational success (Gottfredson, 1997). Arguing for a multidimensional model, Wechsler's (1975) hierarchical perspective allowed for individual interpretation of general and specific abilities (see Tulsky, Saklofske, & Ricker, 2003). Despite considerable factor analytic support for the Wechsler composite scores both in North America (Roid, Prifitera, & Weiss, 1993; Wechsler, 1991), and in international editions (Georgas, Weiss, van de Viver, & Saklofske, 2003), socioeconomic and cultural factors clearly impact child performance on these measures (Georgas et al., 2003). As a result, child characteristics such as personality, motivation, and emotional awareness should be examined in relation to IQ scores (Bowman, Markham, & Roberts, 2002), especially given that IQ accounts for only about half of academic and occupational variance (McGhee, 2002).

Some have challenged the utility of intellectual assessment, arguing that experimental analysis of behavior, using interventions based on curricular or behavioral data, is sufficient to understand and ameliorate any child's learning or behavior problem (e.g., Gresham & Witt, 1997; Reschly & Ysseldyke, 2002). This response-to-intervention (RTI) approach typically ignores or minimizes individual intellectual differences, with proponents suggesting only RTI data are necessary for determining LD eligibility. A "paradigm shift" has been called for in school psychology (Reschly & Ysseldyke, 2002), and school psychologists have been admonished to avoid using intelligence tests (e.g., Fletcher et al., 1994), though apparently this "paradigm shift" does not apply to other psychologists or practitioners who use standardized assessment tools for identification and treatment purposes.

Typically citing early studies that deny the utility of aptitude-treatment interactions (e.g., Cronbach, 1975; Ysseldyke & Sabatino, 1973), advocates of global IQ interpretation argue against intelligence subtest interpretation or profile analysis, claiming the practice is based more on fiction than fact (McDermott, Fantuzzo, & Glutting, 1990, 1992; Watkins, 2000). Global IQ advocates suggest there are no definitive studies that support subtest or factor profile analysis, and that global composites have well-documented predictive validity (Glutting, McDermott, Konold, Snelbaker, & Watkins, 1998; Glutting, Youngstrom, Ward, Ward, & Hale, 1997). The implicit assumption is that all children learn the same way, and only global ability or IQ differentiates individual growth on a standard learning curve. This argument has been generalized to clinical populations (e.g., Glutting et al., 1997), and despite findings that clinical groups often differ from controls on Index or subtest scores, it has been suggested that no score profile is reliable or specific enough to be clearly diagnostic (Glutting et al., 1998).

Although these beliefs have been used to discourage practitioners from using Index or subtest profile interpretation in clinical populations (McDermott, Fantuzzo, & Glutting, 1990, 1992), there is a dearth of empirically valid evidence to support this conclusion. As has been argued elsewhere (Fiorello, Hale, McGrath, Ryan, & Quinn, 2001; Hale & Fiorello, 2004; Hale, Fiorello, Kavanagh, Hoeppner, & Gaither, 2001) the "incremental validity" studies often cited by global IQ advocates cannot be used to determine the relative contributions of FSIQ over Index or subtest profile scores in the prediction of outcomes, due to the multicollinearity of the global and subcomponent measures. Hierarchical regression or analysis of covariance techniques that control for FSIQ or other global score variance before examining the predictive validity of factor or subtest profiles is misleading because the predictors are made up of the same variance. With highly collinear predictors, the order of variable entry into the hierarchical regression equation determines whether predictors will be important or not, as the predictor entered first includes the variable's unique variance as well as its shared variance with other predictors.

Global intellectual deficits are only diagnostic for Borderline Intellectual Functioning and Mental Retardation (DSM-IV-TR, 2000), and even then adaptive behavior deficits must also be established for the latter diagnosis. In LD, where diagnosis is derived from apparent academic failure despite adequate ability, composite IQ scores lose their predictive validity (DSM-IV-TR), leading some to question whether intelligence tests should ever be used (Siegel, 2003). While this extreme view is also fraught with practical and theoretical concerns (Kavale, Holdnack, & Mostert, 2003; Hale, Naglieri, Kaufman, & Kavale, 2003), it highlights that unqualified IQ interpretation can result in poor diagnostic decision-making (Berninger, Hart, Abbott, & Karovsky, 1992). Diagnostic sensitivity and specificity can be achieved by examination of Index and/or subtest profiles for children with neurobiological disorders such as autism, LD, and ADHD (e.g., Mayes & Calhoun, 2003), but this information is lost if one focuses solely on global IQ scores.

While a theoretical stance on the nature of intellectual abilities appears relevant in this discussion, it is not. Any intelligence test score is only an indirect estimate of intellectual ability—it is a measure of intellectual *functioning*—with interpretation confounded by a child's prior educational opportunity, environmental experiences, and neuropsychological processes. For example, vocabulary subtests have been used to assess general intelligence because they tend to have high factor loadings on single factor (*g*) solutions of intelligence tests and are stable in the presence of neuropsychological impairment (Groth-Marnat, Gallagher, Hale, & Kaplan, 2000), yet vocabulary knowledge can be taught and modified with experience or formal instruction (e.g., Hale & Fiorello, 2004). Certainly, vocabulary subtests measure acquisition and expression of lexical-semantic knowledge, assuming *adequate* exposure and language competency, but this caveat makes them partly achievement tests. In addition, intellectual skills are demonstrated within a context of other neuropsychological processes that can confound their measurement (e.g., primary sensory or motor functions affect higher-level cortical tasks). How does one define sensory or motor impairment? Does this only occur when a child fails the gross auditory or vision screening tests, or can this occur at a cortical level, where neuropsychological processes become critical for understanding or expressing oneself? Conceiving of intelligence tests as somehow immune from prior exposure, experiential background, skills training, and other cognitive or neuropsychological processes is clearly overly reductionistic and misleading.

Unlike hierarchical approaches, data partitioning using regression commonality analysis (Pedhazur, 1997) allows for a direct comparison of the unique and shared variance among Index or subtest scores in the prediction of global scores such as FSIQ and other important criterion variables. This technique has been used to argue against FSIQ interpretation for children with LD and ADHD, and children who exhibit significant variability among their Index scores, a substantial portion of the population (Fiorello et al., 2001; Hale et al., 2001). The common occurrence of profile variability has been used to argue it is clinically meaningless (Glutting, McDermott, Watkins, Kush, & Konold, 1997), but contrary to this assumption, these commonality studies clearly show that as subtest or factor variability increases, there is less shared variance among the disparate underlying abilities in both clinical and typical populations, not only in the prediction of FSIQ, but in the prediction of reading, mathematics, and written language achievement as well (Fiorello et al., 2001; Hale et al., 2001).

The Fiorello et al. (2001) and Hale et al. (2001) studies reported above merely provide evidence that the FSIQ can be partitioned into unique and shared predictor (factor and/or subtest) variance using commonality analysis, and these variance components differ for subsamples of the population. Instead of using FSIQ, these data suggest the shared variance among all predictors could serve as an alternative measure of global intellectual functioning, as this variance represents what all predictors have in common. Despite long-held beliefs to the contrary, there is no compelling empirical reason to equate FSIQ scores with single factor solutions or psychometric *g*, as the Wechsler FSIQ is not computed using regression coefficients saved from such analyses, but from scaled scores that have equal weight in its calculation. If FSIQ is composed of more shared than unique predictor variance, then FSIQ interpretation makes good clinical sense. However, if there is a substantial portion of FSIQ variance accounted for by unique predictor variance, or lower level commonalities among two or three predictors, then clinical interpretation should occur below the FSIQ level. Fiorello et al. and Hale et al. do not advocate use of ipsative profile interpretation, nor do they indicate that profile variability has diagnostic specificity for subsamples of children. Instead, the data suggest the obtained variance components have the requisite specificity for clinical interpretation, especially when results are confirmed using multiple data sources over time to establish concurrent, predictive, and treatment validity (e.g., Hale & Fiorello, 2004).

The study presented here was undertaken in an attempt to replicate and extend the Fiorello et al. (2001) and Hale et al. (2001) WISC-III commonality findings using the WISC-IV (Wechsler, 2003) clinical samples (LD, ADHD, and TBI). Communality analysis was used to determine unique and shared Index variance components in predicting FSIQ scores for the clinical samples. Consistent with the Fiorello et al. and Hale et al. studies, we expected FSIQ to be composed of mostly unique and lower-level shared variance among the factor Indices, with little variance accounted for by higher-level commonalities, including the shared variance among the four factors.

METHOD

Participants

The study purpose was to establish the unique and shared WISC-IV Verbal Comprehension (VC), Perceptual Reasoning (PR), Working Memory (WM), and Processing Speed (PS) variance components of the global FSIQ Standard Score (SS) in an attempt to replicate the Fiorello et al. (2001) and Hale et al. (2001) findings using 228 participants diagnosed with LD, ADHD, or TBI (including open- and closed-head injury) selected from the Special Validity Group Studies data reported in the *WISC-IV Technical and Interpretive Manual* (TIM; Wechsler, 2003). Collected concurrently with the WISC-IV standardization data, these group data were obtained to examine WISC-IV clinical utility and specificity. As we were primarily concerned with children of "average" intelligence, children from the LD ($n = 128$), ADHD ($n = 71$), and TBI ($n = 29$) groups were included if they had FSIQ scores between 80 and 120 to ensure extreme scores did not affect study results. This truncated FSIQ range indicates that the study results cannot be generalized to children in these clinical groups who are in the Borderline/Extremely Low (i.e., mentally retarded) or Superior (i.e., gifted) ranges of global intellectual functioning. In contrast to Fiorello et al. (2001) and Hale et al. (2001), the study participants were not chosen on the basis of profile variability.

The LD group, aged 7 to 13 years, consisted of 82 males and 46 females primarily of European-American ($n = 91$), African-American ($n = 24$), and Latino-American ($n = 13$) descent. There were comparable numbers of children represented in the 8 through 12 year old groups (n range 18 to 38). According to TIM criteria, children with LD were discrepant in Reading/Math/Writing ($n = 27$), Reading/Writing ($n = 27$), Reading Only ($n = 46$), and Math Only ($n = 28$) achievement domains. Although still in the average range, the LD group FSIQ was lower than those with ADHD and TBI (see Table 1). The LD group performed comparably on the VC and PR Indices, and the WM SS was their lowest. The ADHD group, which consisted of 47 males and 24 females aged 8 to 13 years, was primarily of European-American descent ($n = 58$), with fewer African-Americans ($n = 5$) and Latino-Americans ($n = 6$) represented. Fairly equal numbers of children were in each age group, except for two 13-year-old children. Although every SS was higher in the ADHD group as compared to the LD group, their PR and PS SS were higher than the TBI group as well. The TBI group consisted of 18 males and 11 females aged 6 to 16 (n range = 2 to 6 in each

Table 1. *Descriptive Statistics for WISC-IV Standardization Special Validity Studies Groups*

	LD		**ADHD**		**TBI**			
WISC-IV SS	***M***	***SD***	***M***	***SD***	***M***	***SD***	***F***	***p***
FSIQ	91.66	8.70	99.06	10.25[a]	95.24	9.83	14.39	<.001
VC	94.19	9.29	99.46	10.72[a]	99.14	9.74[a]	7.84	.001
PR	94.60	10.71	102.28	11.73[ab]	97.62	9.14	9.79	<.001
WM	91.62	10.96	98.11	13.24[a]	99.66	12.18[a]	11.44	<.001
PS	93.08	11.78[b]	94.90	11.48[b]	86.24	14.16	5.44	.005

Note. LD = Learning Disability; ADHD = Attention Deficit/Hyperactivity Disorder; TBI = Traumatic Brain Injury; SS = Standard Score; FSIQ = Full Scale Intelligence Quotient; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; PS = Processing Speed.

[a] greater than LD group.

[b] greater than TBI group.

age group) with 19 being European-American, and 4 each being from the African-American and Latino-American groups. The TBI group had higher VC scores than the LD group, but their PS scores were lower than both the LD and ADHD groups. Further information about the children tested and the inclusion criteria are outlined in TIM Appendix D.

Procedure

For the three clinical groups, multiple forced entry regression equations were computed with VC, PR, WM, and PS factors as predictors of FSIQ to obtain variance components for the FSIQ commonality analyses. After entering the obtained R^2 values into a new database, multiple compute statements were used to calculate unique and shared variance predictor components using standard commonality analysis equations (Pedhazur, 1997) for the four predictors. Nonsignificant variance estimates (below 1%) are not reported, unless they were negative, with the largest negative variance estimate noted for each analysis, regardless of significance.

RESULTS

Reported in Table 2 is the FSIQ commonality analysis for the LD group. As can be seen, the FSIQ for this group is largely composed of unique VC, PR, WM, and PS variance (59.3%). The VC and PR factors each uniquely account for approximately 18% of FSIQ variance, while WM (10%) and PS (13%) unique variance is predictably less, but still significant. Of the commonalities, the largest was between VC and PR (11%), adding credibility to the use of General Ability Index (GAI) as an alternative to FSIQ. The next largest commonalities were among verbal (VC and WM; 7%) and nonverbal (PR and PS; 5.3%) measures, but there was a substantial cross-modality PR and WM (3.8%) commonality, attesting to the relationship between nonverbal reasoning or fluid abilities and executive functions (Denckla, 1996). The three-way C_{VCPRWM} accounted for almost 6% of the FSIQ variance, but no other higher-order commonalities, including the commonality of all factors, exceeded 2% of the FSIQ variance. For the LD group, the FSIQ appears to be composed of largely unique and interpretable shared variance components (89%), not shared variance components among all four factors. Surprisingly, only one small negative commonality, C_{VCPS} (−.01), was found for this heterogeneous group.

Table 2. *FSIQ Commonality Analysis for LD Special Validity Study Sample*

	Proportion of Variance Explained			
	VC	**PR**	**WM**	**PS**
U_{VC}	.180			
U_{PR}		.179		
U_{WM}			.104	
U_{PS}				.130
C_{VCPR}	.112	.112		
C_{VCWM}	.070		.070	
C_{PRWM}		.038	.038	
C_{PRPS}		.053		.053
C_{WMPS}			.030	.030
C_{VCPRWM}	.059	.059	.059	
C_{VCPRPS}	.010	.010		.010
C_{PRWMPS}		.018	.018	.018
$C_{VCPRWMPS}$ (*g*)	.017	.017	.017	.017
Unique	.180	.179	.104	.130
Common	.263	.307	.237	.123
Total	.443	.486	.341	.253

Note. Commonalities <.01 not displayed. U = unique variance; C = shared variance; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; PS = Processing Speed; FSIQ = Full Scale IQ.

Table 3. *FSIQ Commonality Analysis for ADHD Special Validity Study Sample*

	Proportion of Variance Explained			
	VC	**PR**	**WM**	**PS**
U_{VC}	.145			
U_{PR}		.157		
U_{WM}			.117	
U_{PS}				.078
C_{VCPR}	.124	.124		
C_{VCWM}	.077		.077	
C_{PRWM}		.037	.037	
C_{VCPRWM}	.080	.080	.080	
C_{VCPRPS}	.040	.040		.040
C_{VCWMPS}	.028		.028	.028
$C_{VCPRWMPS}$ (*g*)	.024	.024	.024	.024
Unique	.145	.157	.117	.078
Common	.459	.305	.249	.180
Total	.604	.462	.366	.258

Note. Commonalities <.01 not displayed. U = unique variance; C = shared variance; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; PS = Processing Speed; FSIQ = Full Scale IQ.

Table 3 shows the similar results for the ADHD FSIQ commonality analysis. For this group, 50% of FSIQ variance was uniquely accounted for by the predictors, ranging from 7.8% (PS) to 15.7% (PR). The unique contributions of VC and PR accounted for approximately a third of FSIQ variance. As was the case for the LD group, an appreciable amount of FSIQ variance was accounted for by the C_{VCPR} (12.4%), and approximately 8% was accounted for by the C_{VCWM}. Unlike the LD group C_{PRPS}, the ADHD group had approximately 15% of FSIQ variance accounted for in commonalities among three predictors, including the C_{VCPRWM} (8%), C_{VCPRPS} (4%), and C_{VCWMPS} (2.8%). Again, relatively little (2.4%) of FSIQ variance was accounted for by the shared variance among all factors. No significant negative commonalities were found for this group, with the largest being C_{PRPS} (−.001).

Similar to the LD and ADHD group studies, the TBI group's (see Table 4) FSIQ was composed of mostly unique (36%) or interpretable shared (43%) predictor variance, with the four factor commonality again relatively small (2.7%). In this group, the unique variance estimates were fairly comparable for all factors. Interesting, the typically large C_{VCPR} was largely absent, but there were large verbal ($C_{VCWM} = 15\%$) and nonverbal ($C_{PRPS} = 16.4\%$) commonalities, which could suggest lateralization of lesion for this sample. However, the C_{PRWM} (11.1%) and C_{VCPS} (4%) crossed over the verbal/nonverbal dichotomy. Approximately 1/5 of FSIQ variance was accounted for by the three-way C_{VCPRWM}, and the C_{VCPRPS} was 4.8%. Although difficult to interpret, these higher-level commonalities clearly show complex predictor relationships for this clinical group. Similar to the LD and ADHD groups, these commonality results suggest that FSIQ interpretation is limited for children with TBI. However, several negative commonalities call into question the validity of the results, including the C_{WMPS} (−.04), C_{VCWMPS} (−.049), and C_{PRWMPS} (−.033), which may be in part due to the heterogeneity of this open- and closed-head injury sample and the very small sample size.

Table 4. *FSIQ Commonality Analysis for TBI Special Validity Study Sample*

	Proportion of Variance Explained			
	VC	PR	WM	PS
U_{VC}	.109			
U_{PR}		.088		
U_{WM}			.055	
U_{PS}				.108
C_{VCWM}	.150		.150	
C_{VCPS}	.040			.040
C_{PRWM}		.111	.111	
C_{PRPS}		.164		.164
C_{VCPRWM}	.211	.211	.211	
C_{VCPRPS}	.048	.048		.048
$C_{VCPRWMPS}$ (g)	.027	.027	.027	.027
Unique	.109	.088	.055	.108
Common	.434	.535	.377	.157
Total	.543	.623	.432	.265

Note. Commonalities <.01 not displayed. U = unique variance; C = shared variance; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; PS = Processing Speed; FSIQ = Full Scale IQ.

DISCUSSION

Interpretation of individual intellectual test results remains a contentious issue. Many sources for clinical practice describe methods of idiographic interpretation of individual strengths and weaknesses, and discourage the interpretation of global scores when they are composed of divergent cognitive processes (e.g., Flanagan & Kaufman, 2004; Kamphaus, 2001; Sattler, 2001). However, others have argued for nomothetic interpretation at the global intellectual score level, as only these summative scores have the sufficient reliability and validity necessary for individual interpretation (Jensen, 1998; Glutting et al., 1997; McDermott et al., 1990, 1992; Watkins & Canivez, 2004). Despite admonishments to the contrary (e.g., McDermott et al., 1990, 1992), many practicing clinicians report using profile interpretation of intelligence test results below the FSIQ (Pfeiffer, Reddy, Kletzel, Schmelzer, & Boyer, 2000), suggesting objective idiographic interpretation methods must be developed to optimize concurrent and predictive validity and avoid diagnostic and treatment error.

Idiographic interpretation of neuropsychological tests is typically warranted if the measures have adequate sensitivity and specificity (e.g., Fennell & Bauer, 1997; Lezak, 1995; Reynolds, 1997), but intellectual measures are sometimes interpreted as a nomothetic "baseline" of global intelligence, neglecting the shared variance among intellectual and neuropsychological measures. Clearly, as the same brain processes and responds to both intellectual and neuropsychological measures, the question remains as to whether the intellectual test scores have sufficient sensitivity and specificity to warrant interpretation below the global intelligence level, especially for clinical samples. In addition, with more recent tests and revisions, intelligence tests have placed more emphasis on neuropsychological and cognitive processes. Determining whether the structure of intellectual functioning is best represented by global scores, such as the WISC-IV FSIQ, or meaningful cognitive composites, such as the VC, PR, WM, and PS Indices, is necessary to guide both practice and research in clinical populations.

In an attempt to replicate and extend previous studies (e.g., Fiorello et al., 2001; Hale et al., 2001), we sought to examine whether WISC-IV interpretation is best represented by the global FSIQ or Index score level of analysis for children with LD, ADHD, and TBI using regression commonality analysis. Advocated in the TIM and recent clinical texts (e.g., Flanagan & Kaufman, 2004), this level of profile analysis is supported by confirmatory factor analytic findings that the VC, PR, WM, and PS factors provide a better fit of the standardization data than does a single factor model (Wechsler, 2003), consistent with a multifactorial view of intelligence, the predominant perspective of many current intellectual researchers (e.g., Daniel, 1997; Neisser et al., 1996). In this study we found convincing evidence to support idiographic Index interpretation over nomothetic interpretation of a global FSIQ score for LD, ADHD, and TBI populations, replicating our previous findings with clinical populations (Fiorello et al., 2001; Hale et al., 2001).

The LD group results suggest that Index score interpretation is highly appropriate, since the unique contributions of these scores in predicting FSIQ was substantial. In addition, the VC-PR commonality was also significant and considerable, supporting GAI interpretation of overall intellectual functioning may be warranted, consistent with the arguments of Flanagan and Kaufman (2004). Smaller, but still significant, portions of variance were explained by the verbal (VC-WM) and nonverbal (PR-PS) commonalities, suggesting that exploration of verbal versus nonverbal skills may be useful with this population, consistent with hemispheric lateralization findings using the WISC-III (Riccio & Hynd, 2000). A three-way commonality, combining VC, PR, and WM, accounted for a notable amount of variance, consisting of tasks with limited processing speed requirements. A cross-modality commonality, PR and WM, may reflect the relationship between nonverbal reasoning or fluid abilities and executive functions (Denckla, 1996). However, the four-way commonality representing shared variance among all four factors, or general intelligence, accounted for less than 2% of the variance, implying that FSIQ is not particularly meaningful for the LD group.

The ADHD group results were similar to the LD group, as the Index scores, the GAI cluster (VC-PR), and the verbal cluster (VC-WM) all contributed significant and substantial amounts of variance in predicting FSIQ. The PR-WM cluster was again significant, linking fluid reasoning and executive functioning, a relationship critical to explore in neuropsychological identification of the disorder (Barkley, 1997). The more complex three-way Index commonalities found in this sample indicate factorial complexity. The VC-PR-WM commonality may reflect the executive requirements of higher level processing, or the common demands for memory processes, such as rote auditory memory, working memory, or long-term memory encoding, storage, and retrieval. In contrast, the VC-PR-PS commonality consists of tasks that do not require as much working memory because the materials remain present throughout item administration or the items can be repeated at the examinee's request. The VC-WM-PS cluster accounted for little variance, and is not readily interpretable in terms of neuropsychological demands or content, but could represent the necessary cortical tone for processing information or executive requirements necessary for retrieving information from long-term memory. The shared variance among all four predictors was again weak for this group, explaining less than 3% of the FSIQ variance. Together, these FSIQ commonality results for the most common high incidence disability groups, children with LD and ADHD, clearly support idiographic interpretation of factor Index scores over global FSIQ.

The TBI group FSIQ commonality results must be interpreted with caution, due to the small sample size, heterogeneity of the sample, and significant negative commonalities. Nonetheless, the FSIQ commonality results were entirely consistent with the other clinical groups, suggesting that idiographic interpretation of the VC, PR, WM, and PS Index scores appears to be warranted. The shared variance among all predictors was much less than the unique and interpretable shared variance among individual Index scores, consistent with the LD and ADHD groups, as well as other clinical samples (e.g., Fiorello et al., 2001; Hale et al., 2001). Other factor clusters are difficult to interpret given the caveats presented above; however, these clusters could suggest complex relationships among predictors, and negative commonalities likely attest the heterogeneity of this open- and closed-head injury sample.

The commonality findings suggest interpretation at the FSIQ level may obscure important aspects of cognitive functioning for children in clinical groups. Idiographic interpretation at the Index score level appears particularly robust for children with LD and ADHD, as these scores are generally reliable and stable (e.g., Flanagan & Kaufman, 2004), and provided significant and substantial portions of FSIQ variance in this study. Other clusters may be worth exploring to generate hypotheses regarding cognitive strengths and weaknesses as part of a complete clinical evaluation that can confirm or refute them (e.g., Hale & Fiorello, 2004)—most notably the VC-PR cluster, a measure of general reasoning or higher level intellectual processes, and the verbal (VC-WM) and nonverbal (PR-PS) clusters, which could have implications for lateralization of function (Riccio & Hynd, 2000). Several less common associations among clusters that cross stimulus-response modalities may reflect underlying neuropsychological processes that are more consistent with our current understanding of hemispheric functions (e.g., Bryan & Hale, 2001; Goldberg, 2001).

LIMITATIONS AND FUTURE RESEARCH

There are several study limitations and needs for future research that must be addressed. First, the commonality results do not extend to children who are very low or high functioning, as the range was restricted to children with FSIQ scores ranging between 80 and 120. The structure of intellectual functioning may be different in these extreme ranges, and future research using children who are gifted or mentally retarded may reveal new relationships among Index predictors of FSIQ and achievement domains. Second, the data for these children were collected as part of the WISC-IV standardization, and met TIM-defined diagnostic criteria for the LD, ADHD, and TBI groups. As a result, different classification methods for group determination may alter results. Another limitation is that there are no concurrent ecological validity data offered, and no data to substantiate that results have treatment validity. Although this cognitive process-intervention association is one of the most elusive in psychological practice, the data presented here suggest efforts to establish such associations may be fruitful if they occur at the single subject level of analysis and are monitored over time to ensure treatment validity. A study strength or weakness, depending on orientation or clinical practice, is the use of Index scores as the level of analysis for these data. These Indices are composed of individual subtests, each with their own unique and shared variances. While more confidence can be placed in an Index level rather than a subtest level of analysis, it is likely that for some children the Indices are composed of disparate subtest scores. These disparate scores likely impact the interpretability of their respective Index scores, similar to the findings for FSIQ presented here. It is worth considering whether negative commonalities or suppressor effects were found in these analyses. Although a few negative commonalities were found, most of these minor suppressor effects would not negate the findings presented here. However, the TBI commonality analysis had several negative commonalities that limit this group's findings. Likely due to small sample size and heterogeneity, these results suggest further study of this diverse population using larger samples and likely separating those with open- and closed-head injury, as well as other causes for brain dysfunction (e.g., neoplasms, cerebral vascular accidents, genetic disorders). Heterogeneity in the LD and ADHD samples was also a likely limitation in these analyses, given what we know about LD (e.g., Rourke, 1994) and ADHD (e.g., Barkley, 1997) subtypes.

IMPLICATIONS

In this article, we present data that suggest the *structure* of cognitive functioning must be examined in clinical groups (LD, TBI, and ADHD), because the global score does not adequately represent the specificity and lack of shared variance among subcomponent parts. Individual Index scores explain more FSIQ variance than any other level of analysis, including the commonality among all four Index scores. This suggests the unique Index-level contributions in explaining Full Scale variance have the specificity necessary for interpretation, especially when supported by hypothesis testing with multiple data sources over time to ensure ecological and treatment validity (Hale & Fiorello, 2004).

The four-way commonality among the VC, PR, WM, and PS Indices, which we interpret as general intelligence, did not contribute enough variance in the prediction of FSIQ to warrant interpretation in most cases, suggesting FSIQ interpretation for children with these disabilities can obscure important diagnostic information about their underlying cognitive processes. In addition, this suggests that using FSIQ to derive "severe discrepancy" measurements for these clinical groups is inappropriate. Instead, using our concordance-discordance model (Hale & Fiorello, 2004), one could identify cognitive/neuropsychological processing strengths and weaknesses, seeking concordance between the processing weaknesses and academic deficits. Identification of these individual differences could lead to individualized interventions designed to meet a child's unique needs. Many claim that there is no association between cognitive abilities and meaningful outcomes, yet we should never accept the null hypothesis in research, and instead explore these relationships at the single subject level of analysis before we dismiss some of the most psychometrically sophisticated clinical tools developed for individual assessment.

Above and beyond the interpretation of Index scores, our results indicate that interpretation of score clusters may yield important information about cognitive functioning. A given score may be conceptualized as being made up of general intelligence, shared abilities, unique ability (specificity), and error variance. At the Index level, the unique ability is most likely a Broad Cognitive Ability in Cattell-Horn-Carroll parlance, while shared variance that crosses Indices may reflect relationships among stimulus properties, underlying neuropsychological processes, or response mode requirements. Index scores are reliable and stable over time (Watkins & Canivez, 2004; Wechsler, 2003), and appear interpretable across a variety of groups (e.g., Flanagan & Kaufman, 2004), suggesting they have sufficient psychometric integrity for interpretation. For a child with a disability such as LD, ADHD, or TBI, the present study replicates previous findings (e.g., Fiorello et al., 2001; Hale et al., 2001) suggesting interpretation of Index scores is warranted and necessary. Cluster exploration beyond this intermediate interpretive level may also be fruitful, yet the unknown reliability of these clusters suggests they are of limited utility unless hypotheses derived from these scores are carefully tested using multiple data sources over time to ensure construct stability (e.g., Hale & Fiorello, 2004). The true test of the utility of FSIQ, Index, or other intellectual assessment scores lies in how accurately they portray an individual's level and pattern of cognitive or neuropsychological functioning. At least for children with disabilities in this study, the global FSIQ does not meet this important test, suggesting practitioners should seldom interpret FSIQ in clinical practice.

Not only do these data support individual idiographic interpretation of Indices or coherent subtest clusters in clinical practice, but they call into question the validity of using global FSIQ in relationship to other measures in neuropsychological research. As the same brain processes and responds to intellectual and neuropsychological tests, these measures should not be considered to be discrete. As has been demonstrated elsewhere (e.g., Fiorello et al., 2001; Hale et al., 2001), we must explore the unique and shared variance among intellectual and neuropsychological predictors, as hierarchical approaches disproportionately represent the value of collinear predictors depending on the order of entry. The first predictors to enter appear disproportionately large because they represent unique and shared predictor variance, while subsequent variables must rely on their unique variance and shared variance not tapped by earlier predictors, making them seem largely irrelevant. For instance, if two highly collinear variables are put in the same regression equation, and beta weights are used to determine order of entry, an insignificant difference in these weights would determine order of entry, making the first variable seemingly important in the prediction, while the second variable will seem less important, or may not even account for a significant amount of dependent variable variance. This fact suggests shared variance among intellectual and neuropsychological tests should be explored when predicting meaningful clinical outcomes, and FSIQ should not be used as a covariate in hierarchical analyses, because these data are not orthogonal, precluding the use of hierarchical approaches in favor of commonality analysis (see Pedhazur, 1997).

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Barkley, R. A. (1997). *ADHD and the nature of self-control.* New York: Guilford.

Berninger, V. M., Hart, T., Abbott, R., & Karovsky, P. (1992). Defining reading and writing disabilities with and without IQ: A flexible developmental perspective. *Learning Disability Quarterly*, *15*, 103–118.

Bowman, D. B., Markham, P. M., & Roberts, R. D. (2002). Expanding the frontier of human cognitive abilities: So much more than (plain) g! *Learning and Individual Differences*, *13*, 127–158.

Bryan, K. L., & Hale, J. B. (2001). Differential effects of left and right hemisphere accidents on language competency. *Journal of the International Neuropsychological Society*, *7*, 655–664.

Cronbach, L. J. (1975). Five decades of public controversy over mental testing. *American Psychologist*, *30*, 1–14.

Daniel, M. H. (1997). Intelligence testing: Status and trends. *American Psychologist*, *52*, 1038–1045.

Denckla, M. B. (1996). Biological correlates of learning and attention: What is relevant to learning disability and attention-deficit hyperactivity disorder? *Journal of Developmental and Behavioral Pediatrics*, *17*, 114–119.

Fennell, E. B., & Bauer, R. M. (1997). Models of inference in evaluating brain-behavior relationships in children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 204–218). New York: Plenum Press.

Fiorello, C. A., Hale, J. B., McGrath, M., Ryan, K., & Quinn, S. (2001). 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*. New York: Wiley.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., Fowler, A. E., & Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology*, *86*, 6–23.

Georgas, J., Weiss, L. W., van de Viver, F. J. R., & Saklofske, D. H. (2003). A cross-cultural analysis of the WISC-III. In J. Georgas, L. W. Weiss, F. J. van de Viver, & D. H. Saklofske (Eds.), *Culture and intelligence: Cross-cultural analysis of the WISC-III* (pp. 277–314). San Diego: Academic Press.

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

Glutting, J. J., Youngstrom, E. A., Ward, T., Ward, S., & Hale, R. L. (1997). Incremental validity of WISC-III factor scores in predicting achievement: What do they tell us? *Psychological Assessment*, *9*, 295–301.

Gresham, F. M., & Witt, J. C. (1997). Utility of intelligence tests for treatment planning, classification, and placement decisions: Recent empirical findings and future directions. *School Psychology Quarterly*, *12*(3), 249–267.

Groth-Marnat, G., Gallagher, R. E., Hale, J. B., & Kaplan, E. (2000). The Wechsler intelligence scales. In G. Groth-Marnat (Ed.), *Neuropsychological assessment in clinical practice* (pp. 129–194). New York: Wiley.

Goldberg, E. (2001). *The executive brain: Frontal lobes and the civilized mind.* New York: Oxford University Press.

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24*, 79–132.

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

Hale, J. B., Fiorello, C. A., Kavanagh, J. A., Hoeppner, J. B., & Gaither, R. A. (2001). WISC-II predictors of academic achievement for children with learning disabilities: Are global and factor scores comparable? *School Psychology Quarterly*, *16*, 31–55.

Hale, J. B., Naglieri, J. A., Kaufman, A. S., & Kavale, K. A. (2003). Specific learning disability classification in the new individuals with disabilities education act: The danger of good ideas. *The School Psychologist*, *58*, 6–13.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 107–129.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger Publishers.

Kamphaus, R. W. (2001). *Clinical assessment of child and adolescent intelligence* (2nd ed.). Needham Heights, MA: Allyn and Bacon.

Kavale, K. A., Holdnack, J. A., & Mostert, M. P. (2003). The feasibility of a responsiveness to intervention approach for the identification of specific learning disability: A psychometric alternative. Paper presented at the National Research Center on Learning Disabilities "Responsiveness to Intervention" Symposium, Kansas City, MO.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). London: Oxford University Press.

Mayes, S. D., & Calhoun, S. L. (2003). Ability profiles in children with autism: Influence of age and IQ. *Autism*, *7*, 65–80.

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

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

McGhee, R. L. (2002). The McGhee prophecies: Commentary on "is g a viable construct for school psychology? *Learning and Individual Differences*, *13*, 197–203.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: explanation and Prediction* (3rd ed.). Orlando, FL: Holt, Rinehart and Winston.

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*(4), 376–385.

Reschly, D. J., & Ysseldyke, J. E. (2002). Paradigm shift: The past is not the future. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (vol. I) (pp. 3–20). Bethesda, MD: National Association of School Psychologists.

Reynolds, C. R. (1997). Measurement and statistical problems in neuropsychological assessment of children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed.) (pp. 180–203). New York: Plenum Press.

Riccio, C. A., & Hynd, G. W. (2000). Measurable biological substrates to Verbal-Performance differences in Wechsler scores. *School Psychology Quarterly*, *15*, 386–399.

Roid, G. H., Prifitera, A., & Weiss, L. G. (1993). Replication of the WISC-III factor structure in an independent sample [Special issue, Monograph, WISC-III series]. *Journal of Psychoeducational Assessment*, *11*, 6–21.

Rourke, B. P. (1994). Neuropsychological assessment of children with learning disabilities: Measurement issues. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 475–514). Baltimore: Paul H. Brookes Pub.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications*. La Mesa, CA: Jerome M. Sattler, Publisher, Inc.

Siegel, L. S. (2003). IQ-discrepancy definitions and the diagnosis of LD: Introduction to the special issue. *Journal of Learning Disabilities*, *36*, 2–3.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, *47.1*, 1–41.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.

Tulsky, D. S., Saklofske, D. H., & Ricker, J. H. (2003). The evolution of the Wechsler scales. In D. S. Tulsky et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III*. San Diego: Academic Press.

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

Watkins, M. W., & Canivez, G. L. (2004). Temporal stability of WISC-III subtest composite: Strengths and weaknesses. *Psychological Assessment*, *16*, 133–138.

Wechsler, D. (1975). Intelligence defined and undefined: A relativistic appraisal. *American Psychologist*, *30*, 135–139.

Wechsler, D. (1991). *The Wechsler intelligence scale for children*—3rd Edition. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003). *The Wechsler intelligence scale for children*-4th Edition. San Antonio, TX: The Psychological Corporation.

Ysseldyke, J. E., & Sabatino, D. A. (1973). Toward validation of the diagnostic-prescriptive model. *Academic Therapy*, *8*, 415–422.