Psychological Assessment
1997, Vol. 9, No. 3, 295-301

1040-3590/97/$3.00

Incremental Efficacy of WISC-III Factor Scores in Predicting Achievement: What Do They Tell Us?

Joseph J. Glutting and Eric A. Youngstrom
University of Delaware

Thomas Ward and Sandra Ward
College of William and Mary

Robert L. Hale
Pennsylvania State University

Four factors underlie the Full Scale IQ (FSIQ) on the Wechsler Intelligence Scales for Children—Third Edition (WISC-III; D. Wechsler, 1991). Scores from these factors possess greater reliability than individual subtest scores, and some argue that factor scores relate to important phenomena of educational and clinical interest, such as academic achievement and classification status. However, the predictive efficacy of factor scores from the WISC-III relative to the Full Scale IQ (FSIQ) has not been clearly established. This study examined the incremental validity of factor scores from the WISC-III in predicting achievement on the Wechsler Individual Achievement Test (WIAT; Psychological Corporation, 1992). Two groups were evaluated: a nationally representative nonreferred sample (n = 283), and a sample of children referred for evaluation (n = 636). Results indicate that while the factors provide a statistically significant increment, the size of this improvement is generally too small to be of clinical significance for either group. In terms of parsimony and efficiency, the FSIQ is the best predictor of different achievement criteria as measured by the WIAT.

Individual intelligence tests, such as the Wechsler Intelligence Scales for Children—Third Edition (WISC-III; Wechsler, 1991), are among the most popular psychological measures given to children and adolescents (Stinnett, Havey, & Oehler-Stinnett, 1994; Wilson & Reschly, 1996). Although they require considerable investment in terms of time and training, IQ tests have an established record of reliability and validity that makes them a benchmark for other psychometric measures. The practical merit and utility of intelligence tests have been debated extensively, and this dialogue has helped clarify the predictive and concurrent validity of general intelligence, as manifested in overall measures of ability such as the Full Scale IQ (FSIQ) from the WISC-III (Board of Scientific Affairs of the American Psychological Association, 1996).

Broadly speaking, the general intelligence measure seems to be the best predictor of academic success; its correlations with measures of future achievement range from .49 to .65 (Board of Scientific Affairs of the American Psychological Association, 1996; Figueroa & Sassenrath, 1989; Kaufman, 1979, 1994; McGrew & Pehl, 1988; Reilly, Drudge, Rosen, Loew, & Fischer, 1985). This association can be even stronger when both ability and achievement are measured concurrently. For instance, Wechsler (1991) has reported correlations ranging from .70 to .81 between the FSIQ and global measures of achievement, such as the overall achievement composite from the Wechsler Individual Achievement Test (WIAT; Psychological Corporation, 1992). The multiple domain scores from the WIAT serve as particularly good criterion measures of achievement because the WIAT was standardized on a large, nationally representative sample of children and demonstrated good construct validity throughout the formal validation process prior to its publication (Psychological Corporation, 1992).

Given the effort involved in completing and scoring the many subtests within a cognitive ability protocol, and knowing the quality of reputation that IQ tests enjoy, it is not surprising that people hope to garner clinically useful information from the instrument beyond a single, global score reflecting general intelligence (e.g., Bannatyne, 1974; Guilford, 1967; Kaufman, 1979, 1994; Sattler, 1992). A variety of interpretive schemes exist, sharing the premise that more discrete measures, such as subtests, factor indexes, or Verbal versus Performance IQs, provide nontrivial information not captured by the FSIQ, information that is pertinent to some important aspect of human growth and development. Given the loss of parsimony in switching from a single, global measure to multiple and more discrete measures, it is incumbent on proponents of strategies such as subtest analysis to demonstrate this sort of incremental validity.

The WISC-III provides four factor scores that are excellent candidates for providing additional information beyond the FSIQ: Verbal Comprehension Index (VCI), Perceptual Organization Index (POI), Freedom From Distractibility Index (FDI), and Processing Speed Index (PSI). Each factor index may reflect different aspects of ability and is derived from separate

Joseph J. Glutting, Department of Educational Studies, University of Delaware; Eric A. Youngstrom, Department of Psychology, University of Delaware; Thomas Ward and Sandra Ward, School of Education, College of William and Mary; Robert L. Hale, Department of Educational and School Psychology and Special Education, Pennsylvania State University.

Correspondence concerning this article should be addressed to Joseph J. Glutting, 211 Willard Hall, Department of Educational Studies, University of Delaware, Newark, Delaware 19716. Electronic mail may be sent via the Internet to glutting@strauss.udel.edu.

subtest scores. Because each index involves more than one correlated subtest, factor scores are more reliable than individual subtests (as per the Spearman-Brown prophecy; Traub, 1991; Wechsler, 1991). Furthermore, factor scores have better reliability, and because they theoretically represent phenomena beyond the sum of measurement error and specificity, they potentially escape some of the pitfalls that beset attempts to interpret profiles of individual subtest scores (Kamphaus, 1993; McDermott, Fantuzzo, & Glutting, 1990; McDermott, Fantuzzo, Glutting, Watkins, & Baggaley, 1992; Watkins & Kush, 1994).

Factor-score interpretation is also consistent with standards for good assessment practice, such as the "top down" hierarchical approach advocated by Kaufman (1979, 1994) and Sattler (1992). Factor scores also are the logical next choice for interpretation after the FSIQ according to Kamphaus's (1993) principle of emphasizing reliable and valid conclusions from testing. Clinical use of the factor scores has become much more convenient now that tables are provided in the administration manual to facilitate computation, and factor scores are prominently featured on the front cover of the WISC–III protocol (Wechsler, 1991).

Factor scores from the WISC–III are thought to capture dimensions of cognitive ability different from those reflected by either the FSIQ or the various subtest scores (see Kaufman, 1994, for a review; Reynolds & Kaufman, 1990). Therefore, it is possible that ability constructs measured by factor deviation quotients might show a stronger association with specific achievement domains. For example, measures of verbal ability, such as the VIQ or VCI factor from the WISC–III, usually correlate more highly with overall measures of achievement than does general intelligence (Figueroa & Sassenrath, 1989; Wechsler, 1991). Likewise, in studies exploring the concurrent validity of the WISC–III, the VCI showed higher average correlations with measures of overall achievements as well as with more verbally oriented achievement criteria (Wechsler, 1991). Conversely, the POI generally showed the weakest association with achievement. The PSI was less related to achievement than both the FDI and VCI, although in one study (Wechsler, 1991) it showed the highest correlation (r = .73) of any factor with the Math score from the Wide Range Achievement Test–Revised (WRAT–R; Jastak & Wilkinson, 1984). There is also some evidence suggesting that the PSI may relate to attentional problems associated with information processing deficits such as those manifest in attention-deficit hyperactivity disorder (ADHD; Schwean, Saklofske, Yackulic, & Quinn, 1993; Worland, North-Jones, & Stern, 1973). Finally, the FDI may be the most promising factor to provide information about achievement above and beyond the FSIQ. The FDI demonstrated a stronger association with individual measures of achievement than did either the POI or PSI during validation of the WISC–III (Wechsler, 1991).

Factor scores thus appear to promise clinical benefits of differentiation among dimensions of cognitive ability while avoiding the pitfalls of method variance plaguing the more common practice of subtest analysis. Despite these psychometric advantages, the utility of factor scores has not been well researched. What remains is to demonstrate the criterion-related validity of the factor scores and to determine whether these more discrete scores provide substantial improvement in prediction of achievement criteria above and beyond the contributions of a global measure of cognitive ability.

In addition to the general question of whether the factor scores, or the VIQ and PIQ, help to improve the prediction of achievement above levels provided by general intelligence, it also is important to determine whether these different ability scores relate to achievement in the same way for referred as well as normal samples of children. Referred children typically exhibit lower average scores on measures of ability and achievement (Kaufman, 1979, 1994; Wechsler, 1991). Kaufman (1994) presented data that many exceptional groups, such as learning disabled, reading disordered, dyslexic, ADHD, and hearing impaired children, on average show POI scores that are more than 10 points higher than the average of their PSI and FDI scores. Thus, it remains to be determined whether associations between measures of intelligence and achievement are invariant for different educational classifications.

Method

Participants

Two groups of children participated. The nonreferred sample consisted of those children who had completed the WISC–III, the WIAT, and the Guide to the Assessment of Test Session Behavior (GATSB; Glutting & Oakland, 1991) during the national standardization of the GATSB. These participants included 283 children, with a mean age of 12.3 (SD = 2.5) years. Seventy-six percent of the children were White, 11.0% African American, 10.6% Hispanic, and 2.4% of other ethnic background. This sample was stratified within ± 2% of 1990 U.S. census data on the background characteristics of gender, ethnicity (White, African American, Hispanic, or other), and parent education level, as well as overall IQs (FSIQs) from the WISC–III (M = 100, SD = 15).

The referred sample comprised 636 children from the eastern region of Virginia. All had been referred by their schools for psychoeducational assessments. The average age was 11.5 years (SD = 2.2); 430 (67.6%) of the children were male; 66.8% were White, 29.1% African American, 2.6% Latin American, and 1.5% of other background. Postevaluation classifications included 55.3% with learning disabilities; 12.2% with emotional disturbances; 3.3% with mental retardation; and 12.6% with other disabilities, including orthopedic impairments, autism, speech/language impairments, hearing impairments, or other health impairments. The remainder of the children (17.8%) were found to be ineligible for services. Classifications were made by multidisciplinary teams on the basis of criteria established by the school division that are consistent with state guidelines in Virginia.

Predictors

All children in both samples completed subtests from the WISC–III necessary to obtain the four factor index scores. The scores on 10 of the subtests (all except Symbol Search, Digit Span, and Mazes) can be used to create a single summary index, the Full Scale IQ. The FSIQ is the best approximation of general cognitive ability, or psychometric g, thought to underlie measures of intelligence. The same 10 subtests also form the conventional Verbal and Performance scales: the Information, Similarities, Arithmetic, Vocabulary, and Comprehension subtests constitute the Verbal IQ, and the Picture Completion, Coding, Picture Arrangement, Block Design, and Object Assembly subtests make up the Performance IQ.

With the addition of two other subtests, the WISC–III can also generate four different factor scores. The Information, Similarities, Vocabulary, and Comprehension subtests form the basis for the Verbal Compre-

hension Index. Picture Completion, Picture Arrangement, Block Design, and Object Assembly are the subtests underlying the Perceptual Organization Index. The Arithmetic and the optional Digit Span subtest combine to form the Freedom From Distractibility Index score, and the Coding and optional Symbol Search subtests are indicators for the proposed fourth factor—the Processing Speed Index. The FSIQ, VIQ, PIQ, and each factor index are expressed as standard scores, with population means of 100 and standard deviations of 15. The VCI and POI are the two dominant factors measured by the WISC–III subtests (accounting for roughly 45% of the variance in subtest scores) and they are highly correlated with the Verbal and Performance IQs (average VIQ–VCI r = .98, PIQ–POI r = .96; Wechsler, 1991). The FDI and PSI appear to represent different but related dimensions of ability, given their substantially lower correlations with the VIQ, PIQ, and other factor index (average rs ranging from .41 to .70 in the standardization sample of the WISC–III).

Criteria

All children participating in the study completed the WIAT. The WIAT was normed on children ranging in age from 5 years 0 months to 19 years 11 months. The test contains eight subtests that can be aggregated into four composite scores: (a) Reading Composite, derived from the Basic Reading and Reading Comprehension subtests; (b) Mathematics Composite, based on Numerical Operations and Mathematics Reasoning subtests; (c) Writing Composite, composed of the Spelling and Written Expression subtests; and (d) Language Composite, formed from the Listening Comprehension and Oral Expression subtests. The composites are all expressed as standard scores with means of 100 and standard deviations of 15. All children in the nonreferred sample completed all eight subtests, as did 345 children in the referred sample. The remaining children in the referred sample had information available for all WIAT composites except for the Language Composite.

Procedure

Examiners conducting the assessments were all experienced in the individual administration of ability and achievement tests. The relative contributions of different factor scores on the WISC–III to prediction of achievement criteria on the WIAT were assessed through a series of hierarchical regressions. Four different achievement scores, the Reading, Mathematics, Writing, and Language composites, each served as the dependent measure in one set of regression analyses. The Full Scale IQ was compared with both the Verbal and Performance IQs as well as the four factor scores (VCI, POI, FDI, and PSI) through block entry and removal within the hierarchical regressions.

Results

Descriptive statistics showed that the nonreferred sample exhibits the expected level and range of scores on the measures (see Table 1). Referred children scored significantly lower on all measures, as expected. Also important to note is the variability of scores in both samples. With the possible exception of the FDI, variation in scores was not significantly different across samples. Observed standard deviations were also generally close to the expected value of 15. These findings indicate that restriction of range is unlikely to have attenuated correlations observed in either sample.

Amount of Improvement in Predicting Achievement Criteria

Block entry multiple regression analyses addressed the main question of interest: Do any of the four factor scores or VIQ or PIQ substantially improve prediction of different achievement criteria above and beyond the contribution made by FSIQ? Table 2 presents the improvement obtained by entering the four factor scores into the model after FSIQ was entered on the first step. Table 2 also presents the unique contribution of each of the five independent variables (FSIQ, VCI, POI, FDI, and PSI) when all are included simultaneously in the regression equation. These values are equivalent to what would be the overall change in model R^2 if the given variable were entered last into the regression, or alternatively they can be thought of as squared part correlations. As a group, the four factors explain an additional 5–16% of the variance in achievement criteria; and in only one case do they provide more than a 9% improvement (while predicting the Writing Composite in the nonreferred sample). The unique contributions of each factor score are also generally small, ranging from approximately 0 to 5% of the variance in the criterion. For Reading, Mathematics, and Writing criteria, the FDI makes the greatest unique contribution of any factor score, ranging from 1.4–5.2% of the variance. Because of the considerable power of the regression analyses, most of the associations remain statistically significant even after making a Bonferroni adjustment to control for Type I error (i.e., only tabled values with one or two asterisks are not significant when compared with a critical value of $p < .0021$ for 24 comparisons to maintain an overall $\alpha = .05$).

Table 1

Descriptive Statistics and Tests for Differences of Means and Variances for the Nonreferred and Referred Samples

Variable	Nonreferred		Referred		Mean	Variance	
	M	*SD*	*M*	*SD*	*t*[a]	*F*	*p*
WISC–III factors							
FSIQ	100.9	14.6	88.5	14.5	12.03*	0.10	.756
VCI	100.6	14.3	90.9	14.9	9.21*	2.30	.129
POI	101.0	14.8	90.6	16.0	9.35*	0.17	.677
FDI	102.3	13.9	88.1	12.2	14.84[b]*	5.07	.025
PSI	103.5	14.3	92.9	16.9	9.17*	1.34	.247
VIQ	100.7	14.4	89.4	14.3	11.00*	0.02	.886
PIQ	101.4	14.6	89.6	15.1	10.99*	0.07	.797
WIAT composites							
Reading	102.0	13.7	83.7	13.3	18.99*	1.07	.302
Math	102.5	14.2	86.0	13.8	16.55*	0.28	.600
Writing	101.1	13.9	81.1	12.7	21.37*	3.82	.051
Language	101.3	14.1	91.2	14.4	8.88[c]*	0.00	.982

Note. n = 283 for the nonreferred sample; n = 636 for the referred sample in all cases except Language Composite, where n = 352. WISC–III = Wechsler Intelligence Scale for Children—Third Edition; WIAT = Wechsler Individual Achievement Test; FSIQ = Full Scale IQ; VCI = Verbal Comprehension Index; POI = Perceptual Organization Index; FDI = Freedom From Distractibility; PSI = Processing Speed Index; VIQ = Verbal IQ; PIQ = Performance IQ.

[a] All t values are based on an equal variance t test with df = 917 unless otherwise specified. [b] Because the variances would be considered unequal if no Bonferroni adjustment were applied, the reported values are based on an unequal variance t test with df = 483.84. [c] Based on an equal variance t test with df = 633.

* $p < .0005$; compare with a Bonferroni-adjusted critical value of $p < .0045$ to maintain overall $\alpha = .05$ for 11 comparisons.

Table 2
The Incremental Contributions of WISC–III Factor Score in Predicting Achievement Criteria on the WIAT

Predictor	% variance explained		% increment[a]	
	Nonreferred	Referred	Nonreferred	Referred
WIAT Reading				
Step 1				
FSIQ	42.1****	36.0****	42.1****	36.0****
Step 2				
Four factors ($df = 4$)[b]	49.7****	44.8****	7.6****	8.8****
FSIQ			0.0	0.0
VCI			0.0	2.0****
POI			0.0	0.0
FDI			1.4*	4.4****
PSI			0.0	0.0
WIAT Math				
Step 1				
FSIQ	55.8****	55.5****	55.8****	55.7****
Step 2				
Four factors ($df = 4$)[b]	65.5****	61.2****	9.8****	5.7****
FSIQ			0.7	0.0
VCI			0.0	0.0
POI			0.0	0.0
FDI			5.2****	3.8****
PSI			0.0	0.0
WIAT Writing				
Step 1				
FSIQ	31.8****	29.8****	31.8****	29.8****
Step 2				
Four factors ($df = 4$)[b]	47.6****	34.9****	15.8****	5.1****
FSIQ			0.0	0.0
VCI			0.0	1.4***
POI			0.0	0.0
FDI			3.5****	2.5****
PSI			2.6***	0.6*
WIAT Language				
Step 1				
FSIQ	28.4****	55.1****	28.4****	55.1****
Step 2				
Four factors ($df = 4$)[b]	33.3****	60.8****	5.0****	5.8****
FSIQ			1.6*	0.0
VCI			0.0	3.0
POI			1.6*	0.0
FDI			0.0	0.0
PSI			0.0	1.1**

Note. WISC–III = Wechsler Intelligence Scale for Children—Third Edition; WIAT = Wechsler Individual Achievement Tests; FSIQ = Full Scale IQ; Four factors = FSIQ four factors, entered as a group; VCI = Verbal Comprehension Index; POI = Perceptual Organization Index; FDI = Freedom From Distractibility Index; PSI = Processing Speed Index.

[a] Unless otherwise indicated, all unique contributions are squared part correlations, equivalent to the change in R^2 if this variable were entered last in a block entry regression. [b] Partialing out FSIQ.

* $p < .05$. ** $p < .005$. *** $p < .0005$. **** $p < .00005$.

The Verbal and Performance IQs add even less to the prediction of achievement criteria than do the factor scores, as evident in Table 3. VIQ and PIQ together accounted for only 2.1–4.6% of the variance above and beyond the FSIQ, and neither the VIQ nor PIQ accounted for significant unique variance in any of the achievement criteria.

Equivalence of IQ–Achievement Relations Across Samples

For the most part, the associations between the various measures of ability and achievement appear strikingly similar across both the nonreferred and referred samples. The possibility that the relations between predictor and criterion differed between the nonreferred and referred samples was assessed by testing the significance of interaction terms between the sample and predictor (as per Cohen & Cohen, 1983). If the regression coefficients for each sample treated separately would be reliably different from each other, then this would manifest as a significant interaction term for the same predictor when the two samples were combined into one general analysis. Given the relatively large size of both samples, the analyses have sufficient power to detect even minute associations among the variables or differences between the groups. Therefore, all results also are framed in terms of effect sizes, and discussed in terms of potential clinical utility.

The possibility that the association between any given ability score and achievement criterion changed across the two samples was explored by testing the interactions between sample (nonreferred vs. referred) and each predictor (as per Cohen & Cohen, 1983). For example, if the degree of relation between FDI and the Math Composite changed between the nonreferred and referred samples, this would be reflected in a significant Sample × FDI interaction term. Even without adjusting critical values to control for Type I error due to the large number of comparisons, very few significant differences were evident between the two groups of children. Only 3 of the 20 possible interactions for the FSIQ or the factor scores were significant when compared with a critical probability of .0025 to adjust for making 20 comparisons. PSI showed a slightly stronger relation to the Writing Composite in the nonreferred sample, $t(1) = 2.27$, $p = .0232$. The FSIQ, VCI, and POI all showed reliable difference in their association with the Language Composite for the nonreferred versus referred children, even after controlling for average differences in achievement, FSIQ, and the other factor scores (all $ps < .0025$). No differences between samples were evident for VIQ or PIQ relating to any of the achievement criteria.

Incidence and Predictive Utility of Unusual and Rare FDI–VCI Discrepancies

The FDI was the factor score that most consistently showed a unique association with the achievement criteria. However, even the FDI showed at most a 5.1% unique contribution to the prediction of any achievement composite. It is difficult to judge whether this amount constitutes a clinically useful increase in prediction. The potential value of identifying both unusual and rare strengths and weaknesses on the FDI was therefore explored further. The VCI was used as a standard for comparison because

Table 3
The Incremental Contributions of Verbal and Performance IQ Scores on the WISC–III in Predicting Achievement Criteria on the WIAT

	% variance explained		% increment[a]	
Predictor	Nonreferred	Referred	Nonreferred	Referred
WIAT Reading				
Step 1				
FSIQ	42.1****	36.0****	42.1****	36.0****
Step 2				
VIQ–PIQ (df = 2)[b]	46.4****	40.5****	4.3****	4.6****
FSIQ			0.0	0.0
VCI			0.0	0.4*
PIQ			0.0	0.0
WIAT Math				
Step 1				
FSIQ	55.8****	55.5****	55.8****	55.5****
Step 2				
VIQ–PIQ (df = 2)[b]	57.9****	58.7****	2.1**	3.2****
FSIQ			0.0	0.0
VIQ			0.0	0.0
PIQ			0.0	0.0
WIAT Writing				
Step 1				
FSIQ	31.8****	29.8****	31.8****	29.8****
Step 2				
VIQ–PIQ (df = 2)[b]	34.6****	32.6****	2.8**	2.8****
FSIQ			0.0	0.0
VIQ			0.0	0.0
PIQ			0.0	0.0
WIAT Language				
Step 1				
FSIQ	28.4****	55.1****	28.4****	55.1****
Step 2				
VIQ–PIQ (df = 2)[b]	31.3****	58.3****	2.9**	3.3****
FSIQ			0.0	0.0
VIQ			0.0	0.0
PIQ			0.0	0.0

Note. WISC–III = Wechsler Intelligence Scale for Children—Third Edition; WIAT = Wechsler Individual Achievement Tests; FSIQ = Full Scale IQ; VIQ = Verbal IQ; PIQ = Performance IQ.
[a] Unless otherwise indicated, all unique contributions are squared part correlations, equivalent to the change in R^2 if this variable were entered last in a block entry regression. [b] Partialing out FSIQ.
* $p < .05$. ** $p < .005$. *** $p < .0005$. **** $p < .00005$.

it is the best proxy for a global measure of intelligence among the four factor scores. On the basis of standard errors of measurement for the factor indexes, an FDI–VCI discrepancy of 13 or more points is sufficiently large that it would only occur by chance 5% of the time. Significant differences need not be rare ones, though (Glutting, McDermott, Watkins, Kush, & Konold, in press); and roughly 95% of the children in the standardization sample showed FDI–VCI discrepancies of 26 points or less.

Chi-square analyses assessed whether children in the referred sample were more likely to show reliable or rare discrepancies between their obtained FDI and VCI scores. In fact, referred children were slightly more likely to show a significant difference, $\chi^2(1) = 5.36$, $p = .0206$. However, more than one third of all the children exhibited a discrepancy of 13 points or more (29.7% of the nonreferred sample vs. 37.6% of the referred sample), and thus a reliable discrepancy is neither sensitive nor specific to referred children. Furthermore, when the prevalence of rare FDI–VCI discrepancies was examined, there were no significant differences between groups: $\chi^2(1) = 0.23$, $p = .6327$; prevalences of 4.6% and 5.4% in the nonreferred and referred samples, respectively.

Although there appear to be no pronounced differences in the prevalence of FDI–VCI discrepancies between referred and nonreferred samples, it is possible that such a marked variation in cognitive abilities would still lead to measurable differences in academic achievement. Eight t tests addressed the question of whether children with significant (13+ point) or rare (27+ point) FDI–VCI splits showed different performance on average on the Reading, Mathematics, Writing, or Language composites from the WIAT. In only one case was there indication of a difference: Children with a rare FDI–VCI split scored 5.9 points lower on the Mathematics Composite, $t(917) = 2.47$, $p = .014$ (but compared with a critical value of $p < .0063$ to maintain overall $\alpha = .05$ for eight comparisons). When children's discrepancies between FDI and VCI were correlated with their performance on the achievement composites, children with VCI > FDI were found to have trivially higher Reading ($r = .10$, $p = .003$), Mathematics ($r = .09$, $p = .007$), and Language ($r = .28$, $p < .0005$) composite scores.

Discussion

The strong correlations observed in this study between a global measure of ability, such as the WISC–III's FSIQ, and various other measures of achievement are consistent with the extensive literature about the concurrent validity of intelligence tests in general (Board of Scientific Affairs of the American Psychological Association, 1996; Kamphaus, 1993; Kaufman, 1994; Wechsler, 1991).

More surprising is the evidence that reference to the Verbal or Performance IQs on average produces negligible increases in the prediction of any achievement criteria in this study. Current findings include the apparent fact that using information about children's VIQs does not help to improve prediction of such verbally related content areas as reading or writing beyond the predictions that would be made on the basis of the same children's Full Scale IQ scores. In other words, in no case did either the VIQ or PIQ provide useful additional information after controlling for the association between FSIQ and the Reading, Math, Writing, or Language composites of the WIAT.

In general, the four factor scores also did not show any substantial increase in prediction of achievement after partialing out FSIQ. The Verbal Comprehension Index never showed an appreciable unique contribution in the nonreferred children, and it showed only a modest (0 to 3%) increment in the referred children. These results may, in part, reflect the fact that the VCI is so highly correlated with the FSIQ that it is difficult for the

VCI to make a distinct contribution as an independent variable (cf. Chatterjee & Price, 1991). The Perceptual Organization Index also demonstrated no substantial and reliable increment with any of the achievement composites. This may reflect the fact that POI usually is less correlated than are other measures of cognitive ability with academic achievement, as well as the fact that POI is largely collinear with the FSIQ.

The Freedom From Distractibility Index shows the most intriguing results. This factor was the subject of considerable investigation on both the Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974) and the WISC–III, and many claims have been made about different associations between FDI and achievement or behavior (cf. Barkley, 1990; Cohen, Becker, & Campbell, 1990; Kaufman, 1979; Wechsler, 1991). The FDI showed modest but significant unique relations to reading, math, and writing as measured by the WIAT composites. Furthermore, the FDI showed the largest part correlations of any cognitive ability score, albeit only uniquely accounting for 1.4–5.2% of the variance in achievement. It is unclear whether these increments represent a clinically useful amount of information. Attempts to explore this matter by identifying children who showed reliable (i.e., statistically significant) and rare (i.e., those with a low prevalence or base rate) FDI–VCI discrepancies showed that reliable splits were not substantially more likely to occur in children referred for psychoeducational evaluation. Likewise, rare FDI–VCI discrepancies were not meaningfully related to achievement levels for either normal or referred children. On the basis of these results, it appears that interpreting the FDI in addition to FSIQ does not provide enough additional information about children's achievement to merit the extra effort.

The Processing Speed Index is the "newest" factor, appearing only with the addition of the Symbol Search subtest to the WISC–III. Thus, the PSI does not have the same history of extensive research as does the "third factor," or FDI. Present findings add little positive information to our knowledge about the criterion validity of the PSI factor. In both nonreferred and referred children, the PSI provided no predictive gains for reading or math achievement. Although modest unique associations might be evident with either writing or language achievement, these too appeared to be so small as to make their clinical utility appear questionable.

This study is limited because it relies solely on the WISC–III as a measure of cognitive ability and on the WIAT as a measure of achievement. Although the Wechsler scales have excellent psychometric qualities and are among the most commonly used instruments, these results would benefit from cross-validation using additional well-established measures of ability and achievement. Given the reliability of the factor scores and their logical position as the next level of interpretation beneath the FSIQ in hierarchical approaches, it also would be useful to explore whether factor scores, from the WISC–III or other measures of cognitive ability, relate meaningfully to other criteria of interest besides achievement. The possibility that FDI relates to the information processing or attentional dimensions important in ADHD is but one example of a hypothesis that would benefit from further investigation. It is also important to recognize that all of the associations examined here are among concurrent measures: It would still be useful to explore whether factor scores provide any sort of predictive advantage over the FSIQ with longitudinal data (e.g., Moffitt, Caspi, Harkness, & Silva, 1993). Finally, it must be stressed that these results indicate that the FSIQ is the best predictor of actual achievement performance and not necessarily the best measure of overall cognitive ability per se (cf. Kaufman, 1994; Phelps, Leguori, Nisewaner, & Parker, 1993, for discussions of how hearing impairments, difficulty in speaking or understanding English, and other factors, such as severe language disorder, might render the PIQ or POI a better measure of cognitive capability; Sattler, 1992).

The results of this study also have clinical implications. Kaufman (1994) has recommended not administering the Mazes subtest from the WISC–III because it does not provide enough additional information to justify the effort. Similarly, Kamphaus (1993) has recommended not interpreting the PSI from the WISC–III until more information becomes available about the validity and utility of this particular factor score. Given the negative findings here, in the absence of other evidence about the criterion validity of Mazes or the PSI, it seems that clinicians are justified in abbreviating the WISC–III by omitting Mazes and the new Symbol Search subtest (which does not figure in the computation of FSIQ and serves only as an indicator for PSI).

More broadly, the results of the present analyses indicate that the FSIQ is the most parsimonious and powerful predictor of academic achievement obtainable from the WISC–III. Using the Verbal IQ, Performance IQ, or the factor scores to predict academic achievement, even in specific content areas, leads to more complex models (and more laborious calculations for the practitioner) that provide meager dividends (e.g., Grossman & Johnson, 1982; Mishra, 1983). This relationship appears to hold true in both referred and nonreferred samples of children. In terms of reliability, validity, and straightforward interpretation, it appears that the most global ability score will generally prove to be most useful in predicting concurrent achievement.

References

Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC scaled scores. *Journal of Learning Disabilities, 7,* 272–273.

Barkley, R. A. (1990). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment.* New York: Guilford Press.

Board of Scientific Affairs of the American Psychological Association. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77–101.

Chatterjee, S., & Price, B. (1991). *Regression analysis by example* (2nd ed.). New York: Wiley.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.

Cohen, M., Becker, M., & Campbell, R. (1990). Relationships among four methods of assessment of children with attention deficit–hyperactivity disorder. *Journal of School Psychology, 28,* 189–202.

Figueroa, R. A., & Sassenrath, J. M. (1989). A longitudinal study of the predictive validity of the System of Multicultural Pluralistic Assessment (SOMPA). *Psychology in the Schools, 26,* 1–19.

Glutting, J. J., McDermott, P. A., Watkins, M. M., Kush, J. C., & Konold, T. R. (in press). The base rate problem and its consequences for interpreting children's ability profiles. *School Psychology Review.*

Glutting, J. J., & Oakland, T. (1991). *Guide to the assessment of test*

session behaviors (GATSB). San Antonio, TX: Psychological Corporation.

Grossman, F. M., & Johnson, K. M. (1982). WISC–R factor scores as predictors of WRAT performance: A multivariate analysis. *Psychology in the Schools, 19,* 465–468.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Jastak, S., & Wilkinson, G. S. (1984). *Wide Range Achievement Test–Revised.* Wilmington, DE: Jastak.

Kamphaus, R. W. (1993). *Clinical assessment of children's intelligence.* Boston: Allyn & Bacon.

Kaufman, A. (1979). *Intelligent testing with the WISC–R.* New York: Wiley.

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

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., Watkins, M. W., & Baggaley, A. R. (1992). Illusions of meaning in the ipsative assessment of children's ability. *The Journal of Special Education, 25,* 504–526.

McGrew, K. S., & Pehl, J. (1988). Prediction of future achievement by the Woodcock-Johnson Psychoeducational Battery and the WISC–R. *Journal of School Psychology, 26,* 275–281.

Mishra, S. P. (1983). Validity of WISC–R IQs and factor scores in predicting achievement for Mexican-American children. *Psychology in the Schools, 20,* 150–154.

Moffitt, T. E., Caspi, A., Harkness, A. R., & Silva, P. A. (1993). The natural history of change in intellectual performance: Who changes? How much? Is it meaningful? *Journal of Child Psychology and Psychiatry, 14,* 455–506.

Phelps, L., Leguori, S., Nisewaner, K., & Parker, M. (1993). Practical interpretations of the WISC–III with language-disordered children. *Journal of Psychoeducational Assessment WISC–III Monograph,* 71–76.

Psychological Corporation. (1992). *Wechsler Individual Achievement Test manual.* San Antonio, TX: Author.

Reilly, T. P., Drudge, O. W., Rosen, J. C., Loew, D. E., & Fischer, M. (1985). Concurrent and predictive validity of the WISC–R, McCarthy Scales, Woodcock-Johnson, and academic achievement. *Psychology in the Schools, 22,* 380–382.

Reynolds, C. R., & Kaufman, A. S. (1990). Assessment of children's intelligence with the Wechsler Intelligence Scale for Children—Revised (WISC–R). In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 127–165). New York: Guilford Press.

Sattler, J. (1992). *Assessment of children* (3rd ed.). San Diego: Author.

Schwean, V. L., Saklofske, D. H., Yackulic, R. A., & Quinn, D. (1993). WISC–III performance of ADHD children. *Journal of Psychoeducational Assessment WISC–III Monograph,* 56–70.

Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12,* 331–350.

Traub, R. E. (1991). *Reliability for the social sciences.* Newbury Park, CA: Sage.

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

Wilson, M. S., & Reschly, D. J. (1996). Assessment in school psychology training and practice. *School Psychology Review, 25,* 9–23.

Worland, J., North-Jones, M., & Stern, J. A. (1973). Performance and activity of hyperactive and normal boys as a function of distraction and reward. *Journal of Abnormal Child Psychology, 1,* 363–377.

Received September 18, 1996
Revision received February 24, 1997
Accepted February 24, 1997 ■