

Orthogonal Higher Order Structure of the Wechsler Intelligence Scale for Children—Fourth Edition

Marley W. Watkins
Pennsylvania State University

According to J. B. Carroll's (1993) 3-stratum theory, performance on any subtest reflects a mixture of both 2nd-order and 1st-order factors. To disentangle these influences, variance explained by the general factor should be extracted first. The 1st-order factors are then residualized, leaving them orthogonal to the general factor and each other. When these methods were applied to the WISC-IV standardization sample, the general factor accounted for the greatest amount of common (71.3%) and total (38.3%) variance. The largest contribution by a first-order factor was 6.5% of total variance. It was recommended that interpretation of the WISC-IV not discount the strong general factor.

Keywords: WISC-IV, intelligence, factor analysis, individual differences, construct validity

The Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV; Wechsler, 2003a) is the latest version of the Wechsler scales for children. With 10 core and 5 supplemental subtests, the WISC-IV provides four index scores as well as a single overall score, the full scale IQ (FSIQ), that represents general intellectual ability. Although based upon the Wechsler Intelligence Scale for Children—Third Edition (WISC-III; Wechsler, 1991), the WISC-IV dropped three WISC-III subtests, included five new subtests, and abandoned the venerable Verbal and Performance IQ scores in favor of the four factor-based indices. Three of the 10 core subtests were newly created for the WISC-IV and around 35% of the items on five other core subtests were revised. Thus, around 60% of the items in the core WISC-IV subtests are new or revised.

One explicit goal for the WISC-IV revision was to “update the instrument’s theoretical foundations” (Wechsler, 2003b, pp. 8) “to reflect the changes and advances in contemporary intelligence theories” (Zhu & Weiss, 2005, p. 298). To that end, the three-stratum theory of Carroll (1993, 2003) was influential (Flanagan & Kaufman, 2004; Wechsler, 2003b), at least implicitly (Alfonso, Flanagan, & Radwan, 2005), and is the foundation for interpretative recommendations proffered for the WISC-IV (Flanagan & Kaufman, 2004; Keith, Fine, Reynolds, Taub, & Kranzler, 2005).

Carroll (2003) theorized that cognitive abilities exist at three levels or strata:

A first, lower-order stratum comprising some 50 to 60 or more narrow abilities that are linearly independent of each other. . . a second stratum comprising approximately 8 to 10 or more broad abilities, also linearly independent of each other; and a third still higher stratum containing only a single, general intellectual ability commonly termed *g*. (p. 5)

Following this scheme, the WISC-IV captures four broad abilities. This structure was verified by Wechsler (2003b) with exten-

sive exploratory and confirmatory factor analyses. Additional support for four first-order factors was provided by Keith (2005) in a reanalysis of the standardization sample and by Watkins, Wilson, Kotz, Carbone, and Babula (in press) with a sample of students referred for special education.

Inexplicably, Wechsler (2003b) did not conduct a higher order factor analysis to verify and describe the multilevel structure of the WISC-IV. According to McClain (1996), it is a mistake to interpret a second-order factor on the basis of first-order factors. Carretta and Ree (2001) noted that doing so will typically inflate the importance of lower-order factors at the expense of the higher order factor. Carroll (1997) recognized this problem and recommended that the influence of the highest order factor first be explained. Then, individual differences on that factor “must be controlled for or partialled out in it would be desirable to show also that a general factor so identified constitutes a true ability, independent of lower-order factors” (p. 144).

To accomplish this goal, Carroll (1993, 1995, 1997, 2003) and others (Carretta & Ree, 2001; Gustafsson & Snow, 1997; McClain, 1996; Ree, Carretta, & Green, 2003) recommended the Schmid and Leiman (1957) procedure, which transforms the first-order factors so that they are orthogonal to each other and to the second-order factor. Variance explained by the general intelligence factor is extracted first. The first-order factors are then residualized of all the variance present in the general factor (McClain, 1996). Following that recommendation, this study employed the Schmid and Leiman (1957) orthogonalization procedure to elucidate the factor structure of the WISC-IV.

Method

The WISC-IV was standardized on a nationally representative sample of 2,200 children aged 6–16 years closely approximating the 2000 United States Census on sex, 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 ($M = 100$, $SD = 15$) is based on the sum of scores from the 10 core subtests.

Correspondence concerning this article should be addressed to Marley W. Watkins, 125 CEDAR Building, Department of Educational and School Psychology and Special Education, Pennsylvania State University, University Park, PA 16802. E-mail: mww10@psu.edu

The intercorrelation matrix of the 10 core subtests for the 2,200 children in the WISC-IV normative sample (Wechsler, 2003b, p. 51) was subjected to exploratory factor analysis as described by Wechsler (2003b). That is, principal axis extraction of four factors, followed by oblique rotation. Factors were then orthogonalized with the Schmid and Leiman (1957) procedure.

Results

Results are presented in Table 1. The WISC-IV general factor accounted for between 22% and 51% of the variance in the core subtests. The VC factor accounted for an additional 19% to 24% of the variance in the three VC subtests. Beyond *g*, the PR factor accounted for 4% to 9% of the variance in the three PR subtests, the WM factor contributed 11% to 12% of the variance in the two WM subtests, and the PS factor provided 20% to 22% of the variance of its two subtests. As expected, the general factor accounted for the greatest amount of total (38.3%) and common (71.3%) variance. Altogether, the general and first-order factors accounted for 53.7% of the total variance, leaving 46.3% unique variance.

Discussion

Although the first-order factor structure of the WISC-IV was presented by Wechsler (2003b), its hypothesized multilevel structure was not addressed. Interpreting a second-level factor on the basis of first-level factors can be misleading (McClain, 1996) because performance on any subtest reflects a mixture of both general and first-order factors. For example, a group of verbal subtests may be called a measure of verbal comprehension, but a proportion of its nonerror variance is contributed by the general factor. In this case, it is necessary to decompose the variance of the verbal subtests into components separately attributable to verbal comprehension and general intelligence (Gustafsson, 1994). This variance decomposition can be accomplished with the Schmid and Leiman (1957) orthogonalization procedure, which transforms the

first-order factors so that they are orthogonal to each other and to the second-order factor.

When applied to the WISC-IV standardization sample, the Schmid and Leiman (1957) transformation demonstrated that the variance of each WISC-IV subtest can be decomposed into several sizable components. The most important was general intelligence: it accounted for more variance in each of the 10 core WISC-IV subtests than any orthogonal first-order factor. For example, *g* explained 51% of the variance in the Vocabulary subtest whereas the VC factor accounted for an additional 24% of that subtest's variance.

Notably, the unique component (a mixture of specific and error variance) was substantial in all subtests. For example, it accounted for around 25% of the variance in the Vocabulary subtest and 64% of the variance in the Picture Concepts subtest. The combined influence of general intelligence and uniqueness dwarfed the contributions made by the four WISC-IV first-order factors. The first-order factors accounted for 4.1% (PR) to 12.1% (VC) of common variance and 2.2% (PR) to 6.5% (VC) of total variance.

This robust manifestation of general intelligence in the WISC-IV was replicated with a sample of students referred for special education (Watkins et al., in press) and is consistent with previous research with other tests and samples (Jensen, 1998; Lubinski, 2000). Additionally, there have typically been no differences in factor structure found across race or ethnicity and gender groups (Reynolds & Ramsay, 2003).

In terms of incremental predictive validity, the WISC-IV FSIQ has accounted for the bulk of variance (approximately 60%) in both reading and mathematics scores (Glutting, Watkins, Konold, & McDermott, in press). Similar results were reported for the WISC-III (Glutting, Youngstrom, Ward, Ward, & Hale, 1997). The preeminence of the FSIQ is probably explained by the principle of aggregation: "broad abilities account for a relatively small proportion of variance in specific tasks but for a substantial proportion of the variance in scores that are aggregated over several

Table 1

Sources of Variance in the Wechsler Intelligence Scale for Children—Fourth Edition Normative Sample ($N = 2,200$) According to an Orthogonalized Higher Order Factor Model

Subtest	General		VC		PR		WM		PS		h^2	u^2
	<i>b</i>	Var	<i>b</i>	Var	<i>b</i>	Var	<i>b</i>	Var	<i>b</i>	Var		
SI	.71	50	.43	19	.09	01	-.02	00	-.04	00	.69	.31
VO	.72	51	.49	24	.01	00	.02	00	-.01	00	.75	.25
CO	.61	37	.46	21	-.05	00	.01	00	.05	00	.59	.41
BD	.65	42	.00	00	.29	08	-.01	00	.05	00	.51	.49
PCn	.56	32	.08	01	.20	04	.01	00	.02	00	.36	.64
MR	.68	46	.00	00	.30	09	.04	00	.00	00	.55	.45
DS	.55	30	.00	00	.03	00	.34	12	-.04	00	.42	.58
LN	.63	39	.06	00	.00	00	.33	11	.05	00	.51	.49
CD	.47	22	.01	00	-.01	00	-.02	00	.47	22	.44	.56
SS	.58	34	.00	00	.04	00	.02	00	.45	20	.54	.46
% Total Var		38.3		6.5		2.2		2.3		4.4	53.7	46.3
% Common Var		71.3		12.1		4.1		4.3		8.2		

Note. VC = verbal comprehension factor; PR = perceptual reasoning factor; WM = working memory factor; PS = processing speed factor; *b* = loading of subtest on factor; Var = percent variance explained in the subtest; h^2 = communality; u^2 = uniqueness; SI = Similarities; VO = Vocabulary; CO = Comprehension; BD = Block Design; PCn = Picture Concepts; MR = Matrix Reasoning; DS = Digit Span; LN = Letter-Number Sequencing; CD = Coding; SS = Symbol Search.

tasks" (Gustafsson & Undheim, 1996, p. 205). Thus, the FSIQ formed by summing over subscale scores is powerfully affected by the *g* factor (Lubinski, 2004). For example, Gustafsson and Undheim (1996) found that 71% of the total variance of the WISC-III FSIQ was due to the *g* factor.

In short, general intelligence accounts for the bulk of the common variance in the WISC-IV and for most of its predictive validity. Given these results, recommendations favoring interpretation of the WISC-IV first-order factor scores over the FSIQ score (Weiss, Saklofske, & Prifitera, 2005; Williams, Weiss, & Rolfhus, 2003) appear to be misguided.

References

- Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 185-202). New York: Guilford Press.
- Carretta, T. R., & Ree, J. J. (2001). Pitfalls of ability research. *International Journal of Selection and Assessment*, 9, 325-335.
- Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge, United Kingdom: Cambridge University Press.
- Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research*, 30, 429-452.
- Carroll, J. B. (1997). Theoretical and technical issues in identifying a factor of general intelligence. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.), *Intelligence, genes, and success: Scientists respond to the bell curve* (pp. 125-156). New York: Springer-Verlag.
- Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 5-21). New York: Pergamon Press.
- Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. Hoboken, NJ: Wiley.
- Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (in press). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAT-II. *Journal of Special Education*.
- Glutting, J. J., Youngstrom, E. A., Ward, T., Ward, S., & Hale, R. (1997). Incremental efficacy of WISC-III factor scores in predicting achievement: What do they tell us? *Psychological Assessment*, 9, 295-301.
- Gustafsson, J.-E. (1994). Hierarchical models of intelligence and educational achievement. In A. Demetriou and A. Efklides (Eds.), *Intelligence, mind, and reasoning: Structure and development* (pp. 45-73). New York: Elsevier.
- Gustafsson, J.-E., & Snow, R. E. (1997). Ability profiles. In R. F. Dillon (Ed.), *Handbook on testing* (pp. 107-135). Westport, CT: Greenwood 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.
- Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger.
- Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 581-614). New York: Guilford Press.
- Keith, T. Z., Fine, J. G., Reynolds, M. R., Taub, G. E., & Kranzler, J. H. (2005, March). *Hierarchical confirmatory factor structure of the WISC-IV*. Paper presented at the annual meeting of the National Association of School Psychologists, Atlanta, GA.
- 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.
- Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904). "'General Intelligence,' objectively determined and measured." *Journal of Personality and Social Psychology*, 86, 96-111.
- McClain, A. L. (1996). Hierarchical analytic methods that yield different perspectives on dynamics: Aids to interpretation. *Advances in Social Science Methodology*, 4, 229-240.
- Ree, M. J., Carretta, T. R., & Green, M. T. (2003). The ubiquitous role of *g* in training. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 262-274). New York: Pergamon Press.
- Reynolds, C. R., & Ramsay, M. C. (2003). Bias in psychological assessment: An empirical review and recommendations. In J. R. Graham & J. A. Naglieri (Eds.), *Handbook of psychology: Assessment psychology* (Vol. 10, pp. 67-93). Hoboken, NJ: Wiley.
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53-61.
- Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., & Babula, T. (in press). Factor structure of the WISC-IV among referred students. *Educational and Psychological Measurement*.
- Wechsler, D. (1991). Wechsler Intelligence Scale for Children—Third Edition. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2003a). Wechsler Intelligence Scale for Children—Fourth Edition. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2003b). Wechsler Intelligence Scale for Children—Fourth Edition: Technical and interpretive manual. San Antonio, TX: Psychological Corporation.
- Weiss, L. G., Saklofske, D. H., & Prifitera, A. (2005). Interpreting the WISC-IV index scores. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 71-100). New York: Elsevier Academic Press.
- Williams, P. E., Weiss, L. G., & Rolfhus, E. L. (2003). *Clinical validity* (WISC-IV Tech. Rep. No. 3). San Antonio, TX: Psychological Corporation.
- Zhu, J., & Weiss, L. (2005). The Wechsler Scales. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 297-324). New York: Guilford Press.

Received February 8, 2005

Revision received June 17, 2005

Accepted July 20, 2005 ■