School Psychology Quarterly
2008, Vol. 23, No. 4, 533–541

1045-3830/08/$12.00 DOI: 10.1037/a0012884

Orthogonal Higher Order Factor Structure of the Stanford-Binet Intelligence Scales—Fifth Edition for Children and Adolescents

Gary L. Canivez
Eastern Illinois University

Orthogonal higher-order factor structure of the Stanford-Binet Intelligence Scales-Fifth Edition (SB-5; Roid, 2003a) for child and adolescent samples is reported. Multiple criteria for factor extraction unanimously supported extraction of only one dimension and a unidimensional model. However, following results from DiStefano and Dombrowski (2006) and theoretical consideration, two factors were extracted and obliquely rotated and further subjected to the Schmid and Leiman (1957) procedure using MacOrtho (Watkins, 2004). Results showed that the largest portions of total and common variance were accounted for by the second-order, global (‘*g*’) factor and interpretation of the SB-5 should focus primarily, if not exclusively, on the general, Full Scale IQ. No evidence for a five-factor solution was found.

Keywords: SB-5, validity, intelligence, hierarchical exploratory factor analysis

Recent revisions of major tests of intelligence such as the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) and the Stanford-Binet Intelligence Scales—Fifth Edition (SB-5; Roid, 2003a) have utilized Carroll’s (1993) model of the structure of cognitive abilities to facilitate subtest and factor selection and to aid in interpretation of scores and performance. Carroll’s (1993, 2003) 3-stratum theory of cognitive abilities is hierarchical in nature and proposes some 50 to 60 narrow abilities at the bottom (Stratum I), 8 to 10 broad ability factors in the middle (Stratum II), and the general ability factor (‘*g*’) at the apex (Stratum III). Because the narrow abilities and broad ability factors are correlated, subtest performance on cognitive abilities tests reflect combinations of both first-order and second-order factors. Because of this Carroll argued that variance from the higher-order factor should be extracted first to residualize the lower-order factors, leaving them orthogonal to each other and the higher-order factor. Thus, variability associated with a higher-order factor is accounted for before interpreting variability associated with lower-order factors. Statistically, this is achieved through the use of the Schmid and Leiman (1957) procedure that was recommended by Carretta and Ree (2001); Carroll (1993, 1995, 1997, 2003); Gustafsson and Snow (1997); McClain (1996); Ree, Carretta, and Green (2003); and Thompson (2004).

To better understand the structure of the WISC-IV, Watkins (2006) applied this approach with the standardization sample of the WISC-IV (Wechsler, 2003). Watkins et al. (2006) also used the Schmid and Leiman (1957) approach with a sample of Pennsylvania students referred for special education evaluations. With the WISC-IV standardization sample, Watkins (2006) found the general (second-order) factor accounted for the greatest amount of total (38.3%) and common (71.3%) variance. The four first-order factors accounted for smaller portions of variance. Watkins et al. (2006) found identical results. Watkins (2006) and Watkins et al. (2006) argued that interpretation of the WISC-IV should focus on the global FSIQ score because of its accounting for most of the common variance and additional

---

Gary L. Canivez is Professor of Psychology at Eastern Illinois University. His research interests include psychometric reliability and validity investigations of measures of intelligence, achievement, psychopathology, and test bias.

The author thanks Marley W. Watkins for comments on an earlier version of this article. Preliminary analyses and portions of these results were presented at the 2007 Annual Convention of the American Psychological Association, San Francisco, CA.

Correspondence concerning this article should be addressed to Gary L. Canivez, Ph.D., Department of Psychology, 600 Lincoln Avenue, Charleston, IL 61920-3099. E-mail: glcanivez@eiu.edu or http://www.ux1.eiu.edu/~glcanivez

research showing its superiority in predictive validity (Glutting, Watkins, Konold, & McDermott, 2006; Glutting et al., 1997).

The SB-5 (Roid, 2003a) is the most recent edition and since the first edition, the concept of general intelligence has been present. However, the SB-5 was a major revision and purports to measure with 10 subtests, not just general intelligence, but also five factors (Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, and Working Memory) within Verbal (five subtests) and Nonverbal (five subtests) domains. Both the SB-5 Examiners Manual (Roid, 2003b) and the SB-5 Technical Manual (Roid, 2003c) present information on the interpretation of the verbal and nonverbal dichotomy (VIQ and NVIQ) in a similar fashion as the WISC-IV (Wechsler, 2003) as well as the interpretation of the five factor scores. In addition, confirmatory factor analysis (CFA) results are presented to support both two factor (Verbal and Nonverbal) and five factor solutions (Roid, 2003c). However, interpretation of test scores requires broad examination and replication of construct validity of those scores and the investigation of the internal structure of the SB-5 is one important method to provide such support.

Roid (2003c) provided information on the structure of the SB-5 using *only* CFA procedures, and argued “using CFA rather than EFA is more scientifically defensible in establishing the construct validity of the SB-5” (p. 108). No exploratory factor analyses (EFA) were reported for the SB-5 (Roid, 2003c). Frazier and Youngstrom (2007) illustrated substantial differences in CFA and EFA results from factor analysis of major intelligence tests that were associated with increases in the number of dimensions claimed to be measured in later versions. If a test showed agreement between EFA and CFA results, this would lead to greater confidence in the latent structure, but if EFA and CFA results do not agree, this would be cause for concern (Gorsuch, 1983).

Recognizing the absence of EFA in the SB-5, DiStefano and Dombrowski (2006) utilized the SB-5 standardization data to independently conduct EFA *and* CFA to test the SB-5 theoretical model proposed by Roid (2003a) partially because only CFA was reported (Roid, 2003c). They also used as the basis for analyses the 20 split halves of the SB-5 standardization data as did Roid (2003c). This was a clever way to create more than two estimators of each of the five purported Stratum II (Carroll, 1993) dimensions measured by the SB-5; although clinicians use only the 10 actual subtest scores as ability estimates. This was particularly important for utilization of CFA procedures where a minimum of three estimators per latent dimension is recommended (Klien, 2005).

DiStefano and Dombrowski (2006) applied multiple criteria (eigenvalues >1, scree test, parallel analysis (PA), and minimum average partial [MAP]) for determining the correct number of factors to extract and retain (Cattell, 1966; Horn, 1965; Lautenschlager, 1989; O’Connor, 2000; Thompson, 2004; Velicer, 1976; Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986) as recommended by Gorsuch (1983). Results of these methods suggested that the SB-5 was measuring one factor for most age groups but the possibility of two factors (Verbal and Nonverbal) may be present for the two youngest age groups.

However, EFA indicated the two factors (Verbal and Nonverbal) were moderately to highly correlated (.61-.76) whereas CFA yielded very high factor correlations (.89-.98); illustrating the presence of a higher-order factor (Gorsuch, 1983; Thompson, 2004) that should be further explored. There was no evidence to support a five-factor model in EFA *or* CFA (four and five factor CFA models produced estimation problems because of the very high factor correlations ranging from .89-.98) as proposed by Roid (2003c). However, no model estimation problems in CFA procedures were reported in the SB-5 Technical Manual. DiStefano and Dombrowski concluded that the SB-5 was a strong measure of general intelligence across all ages with some support for a verbal and nonverbal dichotomy only for 2 to 10 year-olds.

Because the verbal and nonverbal dimensions of the SB-5 were moderately (EFA) to highly (CFA) correlated (DiStefano & Dombrowski, 2006; Roid, 2003c), it is necessary to investigate the higher-order factor structure of the SB-5. DiStefano and Dombrowski (2006) did not examine a hierarchical model for the SB-5, probably because their conclusion was that EFA and CFA results suggested a unidimensional model. Given the influence of Carroll (1993) on the construction of the SB-5, and its interpretive framework, examination of the hierarchical structure of the SB-5 is important. In fact, Roid (2003c) suggested that in independent EFA of the SB-5, "*g* may need to be extracted from the correlations before the residual variance is analyzed for the lower-order factors" (p. 109). Further, when considering a hierarchical model for SB-5 structure the proportion of SB-5 variance associated with the general factor (Stratum III) and the proportion of SB-5 variance remaining at the broad ability level (Stratum II) must be made available for clinicians to adequately determine which, if any, Stratum II dimensions should be interpreted.

The present study utilized data from the three child and adolescent standardization samples 10 subtest correlation matrices published in the SB-5 Technical Manual (Roid, 2003c) to examine the hierarchical factor structure. Analyses were restricted to the 10 subtest correlation matrices in part because correlation matrices for the 20 split halves are not published in the SB-5 Technical Manual (Roid, 2003c) but also because clinicians use only the 10 subtests, not the 20 split-halves of subtests. Portions of SB-5 variance attributed to the first-order and second-order factors using the Schmid and Leiman (1957) procedure as conducted in the Watkins (2006) and Watkins et al. (2006) studies of the WISC-IV were examined. If multiple factors of the SB-5 are to be interpreted, it is imperative to know how variability is apportioned across the first- and second-order factors.

Method

Participants

Participants were members of the three youngest age groups from the SB-5 standardization sample and included 1,400 2 to 5 year olds, 1,000 6 to 10 year olds, and 1,200 11 to 16 year olds. Demographic characteristics are provided in detail in the SB-5 Technical Manual (Roid, 2003c). The standardization sample closely matched the 1998 United States Census data on key demographic variables of geographic region, parent education level, race/ethnicity, and sex.

Instrument

The Stanford-Binet Intelligence Scales—Fifth Edition (SB-5; Roid, 2003a) includes 10 subtests selected to measure five factors (Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, and Working Memory) within Verbal and Nonverbal domains. A global, Full Scale IQ score is provided in addition to Verbal IQ, Nonverbal IQ, and five composite factor scores and are all based on a mean of 100 and standard deviation of 15.

Procedure

Correlation matrices of the 10 SB-5 subtests for the three youngest age groups in the standardization sample were obtained from the SB-5 Technical Manual (Roid, 2003c) and examined separately through principal axis EFA using SPSS 13.0.0 for Macintosh OSX. Oblique rotations (when extracting more than one factor) provided estimation of subtest associations with correlated factors. Multiple criteria as recommended by Gorsuch (1983) were used to determine the number of factors to retain and included eigenvalues greater than 1 (Guttman, 1954), the scree test (Cattell, 1966), standard error of scree (Zoski & Jurs, 1996), Horn's parallel analysis (HPA; Horn, 1965), Minimum Average Partial (MAP; Velicer, 1976), and theoretical consideration. The scree test was used to visually determine the optimum number of factors to retain and the standard error of scree was used as it was reported to be the most accurate objective scree method (Nasser, Benson, & Wisenbaker, 2002). HPA and MAP analyses are more accurate in determining the number of factors to retain (Frazier & Youngstrom, 2007; Thompson & Daniel, 1996; Zwick & Velicer, 1986) and have been recommended as preferred criteria for factor extraction (Velicer, Eaton, & Fava, 2000).

Following Watkins (2006) and Watkins et al. (2006), the present study limited iterations in first-order factor extraction to two in estimating final communality estimates (Gorsuch, 2003). The correlation matrix of the 10 SB-5 subtests for each of the three youngest age groups (2–5, 6–10, 11–16) were subjected to EFA (principal axis extraction of two factors) and followed by oblique (Promax) rotation. The resulting factors were orthogonalized using the Schmid and Leiman (1957) procedure as programmed in the MacOrtho computer program (Watkins, 2004).

Results

The multiple criteria (eigenvalues >1, scree test, standard error of scree, parallel analyses, and MAP) for determining the number factors to extract and retain indicated that only one factor should be extracted; however, theoretical consideration and results from the DiStefano and Dombrowski (2006) study suggested the possible presence of two highly correlated factors (Verbal and Nonverbal) that should be examined.

Figure 1 illustrates scree analysis results from HPA (Watkins, 2000) for the three age groups. MAP analyses were conducted as programmed for SPSS by O'Connor (2000). MAP analyses also indicated that only one factor should be extracted based on one factor producing the smallest average squared correlation. Standard error of scree results (Watkins, 2007) also found only one nontrivial factor. Theoretical consideration and results from the DiStefano and Dombrowski (2006) study suggested the possible presence of two highly correlated factors (Verbal and Nonverbal) in the 2 to 5 and 6 to 10 age groups, so two factors were extracted and further examined. An attempt to force extraction of five factors from each of the three age groups under investigation was made as it is argued that it is better to overfactor than underfactor (Wood, Tataryn, & Gorsuch, 1996) and resulted in theoretically inconsistent and meaningless subtest associations (loadings) with factors in addition to producing singlet (one subtest) factors, low factor loadings, or factors with no salient loadings. As such, only two factors were extracted and examined with the Schmid and Leiman (1957) procedure.

Schmid-Leiman Orthogonalization

Ages 2 to 5

Results for the 2 to 5-year-old age group of the SB-5 standardization sample are presented in Table 1. The correlation between Factor 1 and Factor 2 from the promax rotation for 2 to 5 year olds was .78. The second-order (general) factor accounted for 41.12% of the total variance and 80.52% of the common variance. The general factor also accounted for between 26% and 50% of individual subtest variability. At the first-order level, Factor I accounted for an additional 5.63% of the total variance and 11.02% of the common variance whereas Factor II accounted for an additional 4.32% of the total variance and 8.46% of the common variance. The first- and second-order factors combined to measure 51.07% of the variance in SB-5 scores resulting in 48.93% unique variance (combination of specific and error variance).

Ages 6 to 10

Results for the 6 to 10-year-old age group of the SB-5 standardization sample are presented in Table 2. The correlation between Factor 1 and Factor 2 from the promax rotation for 6 to 10 year olds was .83. The second-order (general) factor accounted for 48.35% of the total variance and 86.79% of the common variance. The general factor also accounted for between 37% and 56% of individual subtest variability. At the first-order level, Factor I accounted for an additional 4.18% of the total variance and 7.50% of the common variance whereas Factor II accounted for an additional 3.19% of the total variance and 5.72% of the common variance. The first- and second-order factors combined to measure 55.72% of the variance in SB-5 scores resulting in 44.28% unique variance (combination of specific and error variance).

Ages 11 to 16

Results for the 11 to 16-year-old age group of the SB-5 standardization sample are presented in Table 3. The correlation between Factor 1 and Factor 2 from the promax rotation for 11 to 16 year olds was .85. The second-order (general) factor accounted for 49.51% of the total variance and 88.96% of the common variance. The general factor also accounted for between 40% and 62% of individual subtest variability. At the first-order level, Factor I accounted for an additional 3.15% of the total variance and 5.66% of the common variance while Factor II accounted for an additional 3.00% of the total variance and 5.38% of the common variance. The first- and second-order factors combined to measure 55.65% of the variance in SB-5 scores resulting in 44.35% unique variance (combination of specific and error variance).

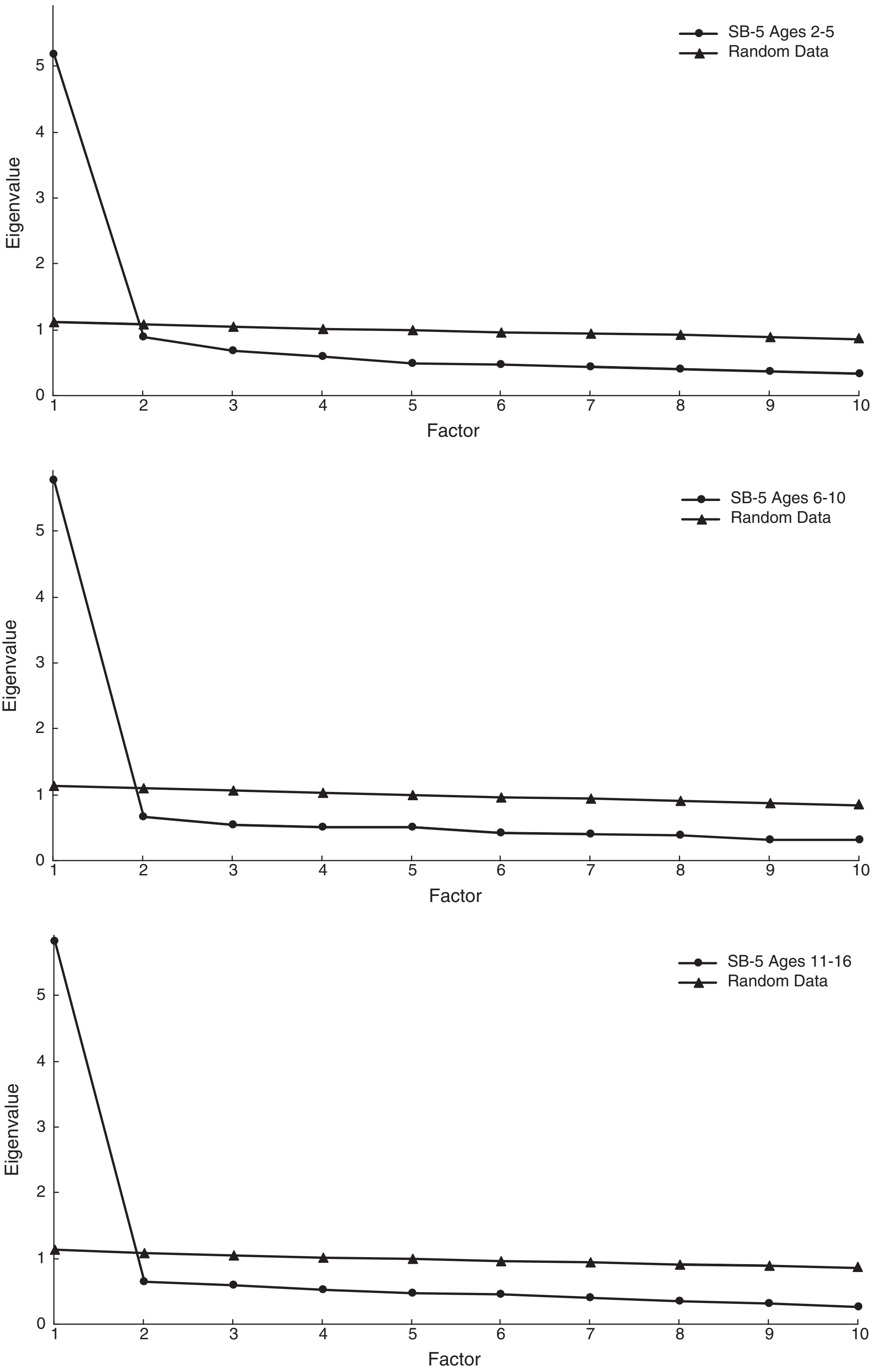

Figure 1. Scree plots of parallel analyses for SB-5 ages 2 to 5, 6 to 10, and 11 to 16 samples.

Table 1
Factor Structure Coefficients and Variance Sources for the SB-5 Standardization Sample Based on the Orthogonalized Higher-Order Factor Model for Ages 2–5 (N = 1,400)

	General		Factor I		Factor II			
Subtest	b	%S^2	b	%S^2	b	%S^2	h^2	u^2
NFR	.51	26	.06	0	.22	5	.32	.68
NKN	.60	36	.38	14	−.06	0	.50	.50
NQR	.71	50	.15	2	.23	5	.58	.42
NVS	.58	33	−.07	0	.38	14	.48	.52
NWM	.65	42	.00	0	.35	12	.54	.46
VFR	.65	43	.23	5	.12	1	.49	.51
VKN	.65	43	.36	13	−.01	0	.56	.44
VQR	.66	43	.19	4	.16	3	.49	.51
VVS	.69	48	.21	4	.17	3	.55	.45
VWM	.69	48	.36	13	.01	0	.61	.39
% Total S^2		41.12		5.63		4.32	51.07	48.93
% Common S^2		80.52		11.02		8.46	—	—

Note. SB-5 = Stanford-Binet Intelligence Scales, Fifth Edition, NFR = Nonverbal Fluid Reasoning, NKN = Nonverbal Knowledge, NQR = Nonverbal Quantitative Reasoning, NVS = Nonverbal Visual-Spatial Processing, NWM = Nonverbal Working Memory, VFR = Verbal Fluid Reasoning, VKN = Verbal Knowledge, VQR = Verbal Quantitative Reasoning, VVS = Verbal Visual-Spatial Processing, VWM = Verbal Working Memory, b = factor structure coefficient (loading), h^2 = communality, u^2 = uniqueness.

Discussion

The present results further illustrate that the SB-5 fundamentally measures general, global intelligence (Stratum III; Carroll, 1993). When examining the 10 SB-5 subtest correlation matrices for the three youngest age groups, there was no evidence to suggest the presence of more than one factor as proposed by Roid (2003c). Factor extraction criteria applied in the present study and in the DiStefano and Dombrowski (2006) study (20 split half subtests) indicated that only one factor should be extracted although the presence of two factors for

Table 2
Factor Structure Coefficients and Variance Sources for the SB-5 Standardization Sample Based on the Orthogonalized Higher-Order Factor Model for Ages 6–10 (N = 1,000)

	General		Factor I		Factor II			
Subtest	b	%S^2	b	%S^2	b	%S^2	h^2	u^2
NFR	.64	41	.23	5	.06	0	.47	.53
NKN	.71	51	.10	1	.23	5	.57	.43
NQR	.72	52	.28	8	.05	0	.60	.40
NVS	.69	48	.27	8	.04	0	.56	.44
NWM	.63	40	.26	7	.03	0	.47	.53
VFR	.71	50	.05	0	.27	7	.57	.43
VKN	.68	47	−.02	0	.33	11	.57	.43
VQR	.75	56	.31	10	.03	0	.65	.35
VVS	.75	56	.13	2	.21	4	.62	.38
VWM	.66	44	.12	1	.18	3	.49	.51
% Total S^2		48.35		4.18		3.19	55.72	44.28
% Common S^2		86.79		7.50		5.72	—	—

Note. SB-5 = Stanford-Binet Intelligence Scales, Fifth Edition, NFR = Nonverbal Fluid Reasoning, NKN = Nonverbal Knowledge, NQR = Nonverbal Quantitative Reasoning, NVS = Nonverbal Visual-Spatial Processing, NWM = Nonverbal Working Memory, VFR = Verbal Fluid Reasoning, VKN = Verbal Knowledge, VQR = Verbal Quantitative Reasoning, VVS = Verbal Visual-Spatial Processing, VWM = Verbal Working Memory, b = factor structure coefficient (loading), h^2 = communality, u^2 = uniqueness.

Table 3
Factor Structure Coefficients and Variance Sources for the SB-5 Standardization Sample Based on the Orthogonalized Higher-Order Factor Model for Ages 11–16 (N = 1,200)

	General		Factor I		Factor II			
Subtest	b	$\%S^2$	b	$\%S^2$	b	$\%S^2$	h^2	u^2
NFR	.61	37	.04	0	.22	5	.42	.58
NKN	.74	55	.26	7	.06	0	.62	.38
NQR	.77	60	.05	0	.28	8	.68	.32
NVS	.64	40	.20	4	.08	1	.45	.55
NWM	.65	43	.04	0	.24	6	.49	.51
VFR	.70	48	.24	6	.05	0	.55	.45
VKN	.69	48	.28	8	.02	0	.56	.44
VQR	.79	62	.09	1	.25	6	.69	.31
VVS	.77	59	.21	5	.12	1	.65	.35
VWM	.65	43	.12	2	.15	2	.46	.54
% Total S^2		49.51		3.15		3.00	55.65	44.35
% Common S^2		88.96		5.66		5.38	—	—

Note. SB-5 = Stanford-Binet Intelligence Scales, Fifth Edition, NFR = Nonverbal Fluid Reasoning, NKN = Nonverbal Knowledge, NQR = Nonverbal Quantitative Reasoning, NVS = Nonverbal Visual-Spatial Processing, NWM = Nonverbal Working Memory, VFR = Verbal Fluid Reasoning, VKN = Verbal Knowledge, VQR = Verbal Quantitative Reasoning, VVS = Verbal Visual-Spatial Processing, VWM = Verbal Working Memory, b = factor structure coefficient (loading), h^2 = communality, u^2 = uniqueness.

the two youngest age groups had some limited support (DiStefano & Dombrowski). No evidence of a five-factor model was found. The unidimensional model seems best supported based on the high factor correlations and high second-order structure coefficients, representing the most parsimonious explanation. As Carretta and Ree (2001) and McClain (1996) pointed out, interpretation of lower-order factors at the expense of higher-order factors is problematic because the subtests contain a mixture of lower-order and higher-order variance. Decomposing variance unique to the lower-order and higher-order components is critical in understanding the structure of the SB-5 and relative importance of interpretive weight to the different scores.

Inspection of orthogonalized factor coefficients for the first-order factors showed some problems of subtest migration (subtest factor coefficients higher on a theoretically different factor), cross-loading (subtests theoretically associated with one dimension/factor were associated with multiple dimensions/factors), and low structure coefficients (see Tables 1-3). Subtest migration problems were noted for the Nonverbal Knowledge subtest for 2 to 5 year olds; Nonverbal Knowledge and Verbal Quantitative Reasoning subtests for 6 to 11 year olds; and Nonverbal Knowledge, Nonverbal Visual-Spatial Processing, and Verbal Quantitative Reasoning subtests for 12 to 16 year olds. Cross loading problems were noted for the Nonverbal Quantitative Reasoning, Verbal Quantitative Reasoning, and Verbal Visual-Spatial Processing subtests for 2 to 5 year olds; Verbal Visual-Spatial Processing and Verbal Working Memory subtests for 6 to 11 and 12–16 year olds. It should also be noted that some of the SB-5 "nonverbal" subtests actually account for more verbal factor variance than nonverbal factor variance; a problem not observed in tests such as the Reynolds Intellectual Assessment Scales (RIAS; Reynolds & Kamphaus, 2003; Nelson, Canivez, Lindstrom, & Hatt, 2007) or WISC-IV (Watkins, 2006; Watkins et al., 2006).

As seen in the orthogonal higher order structure investigation of the WISC-IV by Watkins (2006) and Watkins et al. (2006), and consistent with the observations of Jensen (1998), the present study also found that most of the total and common variance was associated with the general, second-order ('*g*') factor and interpretation of performance on the SB-5 should primarily reside at that level. This was not unexpected given the moderate to high first-order factor correlations, which indicates the presence of a higher-order factor.

On balance, it appears that the SB-5 is a strong measure of general intelligence in children and adolescents but little empirical evidence for additional factors was found. As such, clinicians would be wise to concentrate their interpretation on the overall global IQ score from the SB-5, even with the youngest age groups, where DiStefano and Dombrowski (2006) found some limited evidence for a two-factor (verbal and nonverbal) model. Further, it appears that the SB-5, like many tests of cognitive abilities, may overestimate the number of latent factors (Frazier & Youngstrom, 2007) when disregarding EFA and factor extraction criteria such as HPA and MAP. If assessment and interpretation of Stratum II (Carroll, 1993) factors is of critical importance, test authors will likely need to increase the number of subtests measuring those dimensions to account for greater proportions of variance at that level.

References

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.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245–276.

DiStefano, C., & Dombrowski, S. C. (2006). Investigating the theoretical structure of the Stanford-Binet-Fifth Edition. *Journal of Psychoeducational Assessment, 24,* 123–136.

Frazier, T. W., & Youngstrom, E. A. (2007). Historical increase in the number of factors measured by commercial tests of cognitive ability: Are we overfactoring? *Intelligence, 35,* 169–182.

Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC–IV in estimating reading and math achievement on the WIAI–II. *Journal of Special Education, 40,* 103–114.

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.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Gorsuch, R. L. (2003). Factor analysis. In J. A. Schinka & F. F. Velicer (Eds.), Handbook of psychology: Vol. 2. Research methods in psychology (pp. 143–164). Hoboken, NJ: Wiley.

Gustafsson, J-E., & Snow, R. E. (1997). Ability profiles. In R. F. Dillon (Ed.), *Handbook on testing* (pp. 107–135). Westport, CT: Greenwood Press.

Guttman, L. (1954). Some necessary and sufficient conditions for common factor analysis. *Psychometrica, 19,* 149–161.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185.

Jensen, A. R. (1998). *The g factor.* Westport, CT: Praeger.

Klien, R. B. (2005). *Principles and practice of structural equation modeling.* New York: Guilford Press.

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research, 24,* 365–395.

McClain, A. L. (1996). Hierarchical analytic methods that yield different perspectives on dynamics: Aids to interpretation. *Advances in Social Science Methodology, 4,* 229–240.

Nasser, F., Benson, J., & Wisenbaker, J. (2002). The performance of regression-based variations of the visual scree for determining the number of common factors. *Educational and Psychological Measurement, 62,* 397–419.

Nelson, J. M., Canivez, G. L., Lindstrom, W., & Hatt, C. (2007). Higher-order exploratory factor analysis of the Reynolds Intellectual Assessment Scales with a referred sample. *Journal of School Psychology, 45,* 439–456.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32,* 396–402.

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., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales.* Lutz, FL: Psychological Assessment Resources, Inc.

Roid, G. (2003a). *Stanford-Binet Intelligence Scales: Fifth Edition.* Itasca, IL: Riverside Publishing.

Roid, G. (2003b). *Stanford-Binet Intelligence Scales: Fifth Edition, Examiner's Manual.* Itasca, IL: Riverside Publishing.

Roid, G. (2003c). *Stanford-Binet Intelligence Scales: Fifth Edition, Technical Manual.* Itasca, IL: Riverside Publishing.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22,* 53–61.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC: American Psychological Association.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement, 56,* 197–208.

Velicer, W. F. (1976). Determining the number of components form the matrix of partial correlations. *Psychometrika, 31,* 321–327.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A view and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: A festschrift to Douglas Jackson at seventy* (pp. 41–71). Norwell, MA: Kluwer Academic.

Watkins, M. W. (2000). *Monte Carlo PCA for parallel analysis* [Computer Software]. State College, PA: Author.

Watkins, M. W. (2004). *MacOrtho.* [Computer Software]. State College, PA: Author.

Watkins, M. W. (2006). Orthogonal higher order structure of the Weschsler Intelligence Scale for Children—Fourth Edition. *Psychological Assessment, 18,* 123–125.

Watkins, M. W. (2007). *SEscree* [Computer software]. State College, PA: Ed. & Psych Associates.

Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., & Babula, T. (2006). Factor structure of the Wechsler Intelligence Scale for Children-Fourth Edition among referred students. *Educational and Psychological Measurement, 66,* 975–983.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—Fourth Edition: Technical and interpretive manual.* San Antonio, TX: Psychological Corporation.

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and over-extraction on principal axis factor analysis with varimax rotation. *Psychological Methods, 1,* 354–365.

Zoski, K. W., & Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: The standard error scree. *Educational and Psychological Measurement, 56,* 443–451.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 117,* 253–269.