Orthogonal Higher Order Structure of the WISC-IV Spanish Using Hierarchical Exploratory Factor Analytic Procedures

Ryan J. McGill1 and Gary L. Canivez2

Abstract
As recommended by Carroll, the present study examined the factor structure of the Wechsler Intelligence Scale for Children–Fourth Edition Spanish (WISC-IV Spanish) normative sample using higher order exploratory factor analytic techniques not included in the WISC-IV Spanish Technical Manual. Results indicated that the WISC-IV Spanish subtests were properly aligned with theoretically proposed factors; however, application of the Schmid and Leiman procedure found that the g factor accounted for large portions of total and common variance, whereas the four first-order factors accounted for small portions of total and common variance. Implications for clinical interpretation of the measurement instrument are discussed.

Keywords
WISC-IV Spanish, exploratory factor analysis, intelligence, Schmid–Leiman procedure

The Wechsler Intelligence Scale for Children–Fourth Edition Spanish (WISC-IV Spanish; Wechsler, 2005a) is a translation and adaptation of the WISC-IV (Wechsler, 2003a) for use with Spanish-speaking children and adolescents aged 6 to 16 years. It features 10 core and four supplemental subtests, and provides four index scores as well as an overall full scale IQ score (FSIQ) that represents general intellectual ability. According to Harris, Muñoz, and Llorente (2008), the WISC-IV Spanish was developed to meet the needs of examinees learning English as a second language and in the process of acculturating to the United States. It should be noted that although the WISC-IV was recently revised, the WISC-IV Spanish has yet to be updated and continues to be widely utilized by practitioners (Sotelo-Dynega & Dixon, 2014).

According to the Technical Manual (Wechsler, 2005b), an explicit goal in developing a Spanish language version of the WISC-IV was to create a measurement instrument that explicitly reflected the content, structure, and theoretical foundations of the WISC-IV (attempts to develop a Spanish version of the WISC-III were never completed). To that end, WISC-IV Spanish users are advised in the Technical Manual to become familiar with the rationale for the revisions made to the WISC-IV from the previous iteration of the measurement instrument. According to the

1Texas Woman’s University, Denton, USA
2Eastern Illinois University, Charleston, USA

Corresponding Author:
Ryan J. McGill, Department of Psychology and Philosophy, Texas Woman’s University, P.O. Box 425470, Denton, TX 76204, USA.
Email: rmcgill@twu.edu
Technical Manual (Wechsler, 2003b), a major goal of the WISC-IV revision was to update the instrument’s theoretical foundation to better reflect advances in psychometric theory, including elements of Carroll’s (1993) three-stratum theory of cognitive abilities.

Despite being criticized as atheoretical, subsequent factor analytic research (e.g., Keith, Fine, Taub, Reynolds, & Kranzler, 2006; Weiss, Keith, Zhu, & Chen, 2013) indicated that the WISC-IV measures several constructs consistent with the Cattell–Horn–Carroll (CHC) model of intelligence, including crystallized ability, fluid reasoning, and processing speed (PS). Accordingly, users have been encouraged (e.g., Flanagan & Kaufman, 2009) to adopt a CHC-based interpretive scheme for WISC-IV scores that places primary emphasis on evaluation of the profile of obtained first-order factor scores. As a result of the structural congruence between the measures, these results and interpretive recommendations are thought to be germane for the WISC-IV Spanish (San Miguel Montes, Allen, Puente, & Neblina, 2010).

The structure of the WISC-IV Spanish was examined using extensive exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) supporting four first-order factors consistent with the WISC-IV (Wechsler, 2005b). However, the factor solution initially examined in exploratory analyses appeared to be forced to comport with publisher theory as no empirical support for a four-factor extraction was provided. Inexplicably, higher order factor analysis to verify and describe the hierarchical structure of the WISC-IV Spanish was not conducted despite the implied influence of Carroll’s (1993) higher order model. According to Carroll (1993, 1995), subtest performance on tests of cognitive ability reflects the combination of both first-order and second-order factors. As a result, he argued that variance from the higher order factor must be extracted first to residualize the lower order factors, leaving them orthogonal to the higher order dimension. Variability associated with a higher order factor is accounted for before interpreting variability associated with lower order factors, resulting in variance being apportioned correctly to higher order and lower order dimensions. To accomplish this, Carroll (1995) recommended second-order factor analysis of first-order factor correlations followed by a Schmid–Leiman transformation (Schmid & Leiman, 1957). This process transforms an oblique solution “into an orthogonal solution which not only preserves the desired interpretation characteristics of the oblique solution, but also discloses the hierarchical structuring of the variables” (Schmid & Leiman, 1957, p. 53).

Also missing from the Technical Manual were empirical support for a four-factor extraction, proportions of variance accounted for by the higher order FSIQ and the four first-order factors, higher order subtest loadings, and subtest specificity estimates. As a result, clinicians do not have the information necessary for determining the relative importance of the WISC-IV Spanish factor and subtest scores relative to the FSIQ score, which can result in overinterpretation of lower order factors at the expense of the higher order factor (Carretta & Ree, 2001). Following the recommendations of Carroll (1995), the present study used exploratory factor analytic procedures followed by the Schmid and Leiman (1957) orthogonalization procedure to better clarify the hierarchical factor structure and allocation of reliable score variance in WISC-IV Spanish as illustrated by Watkins (2006) and Canivez, Watkins, and Dombrowski (2015).

**Method**

The WISC-IV Spanish was standardized on a stratified nationally representative sample of 500 Spanish-speaking children aged 6 to 16 years closely approximating the 2001 U.S. Census on age, sex, and parent education level. The WISC-IV Spanish has 10 core subtests (M = 10, SD = 3) that combine to form four-factor composites (M = 100, SD = 15): Verbal Comprehension (VC), Perceptual Reasoning (PR), PS, and Working Memory (WM). The FSIQ (M = 100, SD = 15) is based on the sum of scores from the 10 core subtests.

As recommended by Gorsuch (2003), multiple criteria for determining the number of factors to retain were examined. Consistent with best practice recommendations in the
professional literature (e.g., Frazier & Youngstrom, 2007), these procedures included Horn’s parallel analysis (HPA; Horn, 1965) and minimum average partial (MAP; Velicer, 1976). HPA was generated using the Monte Carlo Principle Component Analysis (PCA) for Parallel Analysis program (Watkins, 2000) with 100 replications to produce stable estimates. MAP procedures were conducted using O’Connor’s (2000) SPSS syntax. The intercorrelation matrix of the 10 core subtests for the WISC-IV Spanish normative sample of 500 children and adolescents (Wechsler, 2005b) was then subjected to EFA as described in the Technical Manual (Wechsler, 2005b), that is, principal-axis extraction, followed by oblique rotation. The resulting factor correlations were subjected to second-order EFA and orthogonalized using the Schmid and Leiman (1957) procedure as programmed in the MacOrtho program (Watkins, 2004). Omega-hierarchical and omega-subscale coefficients (Reise, 2012) were also estimated with the Omega program (Watkins, 2013), which is based on the tutorial by Brunner, Nagy, and Wilhelm (2012) and the work of Zinbarg, Revelle, Yovel, and Li (2005) and Zinbarg, Yovel, Revelle, and McDonald (2006).

Results

Parallel analysis (Horn, 1965) and MAP (Velicer, 1976) criteria suggested that one factor be retained for the WISC-IV Spanish normative sample. In sum, none of the extraction criteria supported the retention of four factors, as suggested by the WISC-IV Spanish Technical Manual, with the exception of publisher theory. Despite factor extraction criteria suggesting extraction of only one factor, four factors were extracted to comport with the proposed test structure indicated in the Technical Manual. It is believed that these methods provide for a direct examination of publisher theory with results that are potentially relevant for clinicians who utilize the WISC-IV Spanish in practice.

Schmid–Leiman variance partitioning EFA results are presented in Table 1. Correlations between the four first-order factors from the promax rotation \((k = 4)\) ranging from .62 to .74 suggested the presence of a higher order dimension requiring additional explication. The WISC-IV Spanish general factor accounted for between 25% and 47% of the variance in the core subtests. The VC factor accounted for an additional 15% to 26% of the variance in the three VC subtests. Beyond \(g\), the PR factor accounted for between 4% and 17% of the variance in the three PR subtests, the PS factor provided 11% to 14% of the variance in the PS subtests, and the WM factor contributed 9% of the variance in its two subtests. As a consequence, the general factor accounted for the greatest amount of total (41.3%) and common variance (73.2%). Altogether, the general and first-order factors accounted for 56.6% of the total variance, leaving 43.4% unique variance (combination of specific and error variance). Omega-hierarchical coefficient for general intelligence (.833) was high, supporting scale interpretation but omega-subscale coefficients for the four WISC-IV Spanish group factors ranged from .120 (WM) to .280 (VC) and likely contained too little true score variance for clinical interpretation (Reise, 2012; Reise, Bonifay, & Haviland, 2013).

Discussion

Although the first-order factor structure of the WISC-IV Spanish was presented in the Technical Manual (Wechsler, 2005b), its hypothesized higher order structure was not investigated. Because first-order factors are abstractions of measured variables, interpreting a second-order factor on the basis of the relationships between these variables can be misleading because performance on any cognitive subtest reflects a mixture of both general and first-order factors (Watkins & Beaujean, 2014). As a result, Carroll (1995) insisted that it is necessary to decompose variance into components that can be sourced more appropriately to higher and lower
order dimensions. Given the fact that the Technical Manual encourages users to interpret the WISC-IV Spanish predominately at the first-order factor level, it is vital that the higher order structure be examined to determine how well the scores measure those traits (i.e., American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014).

The application of the Schmid and Leiman (1957) transformation to the WISC-IV Spanish normative sample demonstrated that variance for each of the core subtests can be decomposed into multiple components. The most important of these components was a higher order latent dimension ($g$). General cognitive ability accounted for more variance in each of the WISC-IV Spanish subtests than any of the proposed first-order factors. Across all subtests, the combination of $g$ and uniqueness overshadowed the contributions made by the four first-order factors (see Table 1).

Notably, although subtest loadings were generally consistent with their theoretically assigned factors, the four-factor model appears to be overfactored (e.g., Frazier & Youngstrom, 2007). Specifically, once the reliable variance attributable to the higher order dimension was partialled out, Letter–Number Sequencing failed to load saliently ($≥ .30$) on the WM factor. These results suggest that WM may not be a viable latent dimension on the WISC-IV Spanish because a common factor cannot be produced from a singlet loading of one manifest variable (Preacher & MacCallum, 2003).

The results from the present investigation are relatively consistent with examinations of the higher order structure of the WISC-IV (e.g., Canivez, 2014; Watkins, 2006). As observed in the present investigation, Canivez (2014) and Watkins (2006) both found that the general factor accounted for the largest portions (> 70%) of common variance in the WISC-IV subtests in both normative and referred samples. Interestingly, both studies provided evidence for the viability of the WM factor. Whereas Canivez (2014) found that the subtests theoretically assigned to the WM factor aligned with that factor when estimating a direct hierarchical model via CFA, Watkins (2006) found that both of those measures loaded saliently on that dimension.

### Table 1. Sources of Variance in the WISC-IV Spanish for Ages 6 to 16 ($N = 500$): 10 Core Subtests According to an Orthogonalized (Schmid & Leiman, 1957) Higher Order Factor Model.

<table>
<thead>
<tr>
<th>Subtest</th>
<th>General</th>
<th>VC</th>
<th>PR</th>
<th>PS</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$</td>
<td>$S^2$</td>
<td>$b$</td>
<td>$S^2$</td>
<td>$b$</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.682</td>
<td>.465</td>
<td>.513</td>
<td>.263</td>
<td>-.024</td>
</tr>
<tr>
<td>Comprehension</td>
<td>.613</td>
<td>.376</td>
<td>.494</td>
<td>.244</td>
<td>-.061</td>
</tr>
<tr>
<td>Similarities</td>
<td>.720</td>
<td>.518</td>
<td>.392</td>
<td>.154</td>
<td>.128</td>
</tr>
<tr>
<td>Matrix reasoning</td>
<td>.718</td>
<td>.516</td>
<td>-.012</td>
<td>.000</td>
<td>.410</td>
</tr>
<tr>
<td>Block design</td>
<td>.661</td>
<td>.437</td>
<td>-.056</td>
<td>.000</td>
<td>.359</td>
</tr>
<tr>
<td>Picture concepts</td>
<td>.658</td>
<td>.433</td>
<td>.160</td>
<td>.026</td>
<td>.208</td>
</tr>
<tr>
<td>Coding</td>
<td>.498</td>
<td>.248</td>
<td>.024</td>
<td>.001</td>
<td>-.022</td>
</tr>
<tr>
<td>Symbol search</td>
<td>.584</td>
<td>.341</td>
<td>-.029</td>
<td>.000</td>
<td>.079</td>
</tr>
<tr>
<td>Digit span</td>
<td>.636</td>
<td>.404</td>
<td>.057</td>
<td>.003</td>
<td>.003</td>
</tr>
<tr>
<td>Letter–number Sequencing</td>
<td>.634</td>
<td>.402</td>
<td>.080</td>
<td>.006</td>
<td>.077</td>
</tr>
</tbody>
</table>

% total variance: 41.3 7.0 3.7 2.7 1.8 56.6 43.4

% common variance: 73.2 12.3 6.5 4.8 3.2 100

$\omega_h$ = Omega Hierarchical (general); $\omega_s$ = Omega Subscale (group factors).

Note. Bold denotes theoretically consistent factor loadings. WISC-IV = Wechsler Intelligence Scale for Children–Fourth edition; VC = Verbal Comprehension Index; PR = Perceptual Reasoning Index; PS = Processing Speed Index; WM = Working Memory Index; $b$ = standardized loading of subtest on factor; $S^2$ = variance explained; $h^2$ = communality; $u^2$ = uniqueness; $\omega_h$ = Omega Hierarchical (general); $\omega_s$ = Omega Subscale (group factors).
after higher order variance partitioning via the Schmid–Leiman procedure in the normative sample. Given the fact that the WISC-IV normative sample \((N = 2,200)\) was significantly larger than the present sample \((N = 500)\), it is certainly possible that the weak WM loadings produced for the WISC-IV Spanish are an artifact of sampling error (MacCallum, Widaman, Preacher, & Hong, 2001).

Given the hierarchical nature of the WISC-IV Spanish, additional research estimating rival measurement models through CFA such as a bifactor structure and computing model-based reliability coefficients (e.g., Reise et al., 2013) would be beneficial. Nevertheless, it is believed that the present results are a first step in examining the higher order structure of the WISC-IV Spanish and will be a useful criterion to compare against future results produced from alternate methods (e.g., CFA). In addition, these results may be instructive for guiding the development of future iterations of the WISC-IV Spanish.

To conclude, our results diverge from the four-factor structure posited in the WISC-IV Spanish Technical Manual and therefore suggest caution in interpretation of first-order factors until additional research has been conducted. As a consequence, it is recommended that users of the WISC-IV Spanish focus most, if not all, of their interpretive weight on the FSIQ as \(g\) accounts for the majority of common variance in the WISC-IV Spanish subtests, and too little true score variance is contained in the group factors. Thus, clinicians who wish to interpret beyond the FSIQ must account for its effects at every level of the WISC-IV Spanish or risk overinterpretation (Glutting, Watkins, Konold, & McDermott, 2006). These results suggest that if assessment and interpretation of first-order dimensions are of critical importance, the test authors will likely need to increase the number of subtests estimating those dimensions to increase the amount of variance at that level of measurement (Canivez, 2011).

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**References**


