Exploring the Latent Structure of the Luria Model for the KABC-II at School Age: Further Insights from Confirmatory Factor Analysis

Ryan J. McGill

College of William & Mary

Author Note

Ryan J. McGill, School of Education, College of William & Mary, P.O. Box 8795, Williamsburg, VA 23187.

Correspondence concerning this article should be addressed to Ryan J. McGill, School of Education, College of William & Mary, P. O. Box 8795 Williamsburg, VA. 23187. E-Mail: rmcgill@wm.edu

Abstract

The present study examined the factor structure of the Luria interpretive model for the Kaufman Assessment Battery for Children-Second Edition (KABC-II; Kaufman & Kaufman, 2004a) with normative sample participants aged 7-18 (*N* = 2,025) using confirmatory factor analysis with maximum likelihood estimation. For the eight subtest Luria configuration, an alternative higher-order model with Pattern Reasoning being permitted to cross-load on the Planning and Simultaneous Processing factors (as per McGill & Spurgin, 2015 and Reynolds et al., 2007) provided the best fit to the normative sample data. Variance apportionment suggests that additional consideration, beyond the omnibus MPI, of the contribution of the first-order factor-based scores (i.e., SQ, SM, P and L), and in some cases the individual subtests themselves, may be warranted. Implications for clinical interpretation and the anticipated normative update of the measurement instrument are discussed.

Keywords: Luria model, KABC-II, Structural validity

Exploring the Latent Structure of the Luria Model for the KABC-II at School Age: Further Insights from Confirmatory Factor Analysis

The Kaufman Assessment Battery for Children-Second Edition (KABC-II; Kaufman & Kaufman, 2004b) is an individually administered battery of cognitive tests for children and adolescents ages 3-18 years. The KABC-II was developed to reflect core elements of both Luria's (1973) neuropsychological theory of processing and the Cattell-Horn-Carroll (CHC; Schneider & McGrew, 2012) psychometric model of broad and narrow cognitive abilities. Due to its clinical and theoretical flexibility, the instrument is widely utilized by school psychologists (Sotelo & Dynega-Dixon, 2014) and has served as a reference tool for researchers seeking to further our understanding of the latent structure of cognitive abilities as well as cognitive-achievement relations (e.g., Benson, Kranzler, & Floyd, 2016; Kaufman et al., 2012; Reynolds, Keith, Flanagan, & Alfonso, 2013).

Examiners may elect to interpret the KABC-II using either a CHC- or Luria-based interpretive model though they must decide *a priori* which scheme to use prior to the beginning of testing. This choice is not arbitrary as the interpretive models differ both in terms of factor structure (four first-order factors in the Luria model versus five factors in the CHC model) and whether indicators of acquired knowledge (i.e., Crystallized Ability [Gc]) are administered to examinees, as it is noted in the manual that Luria considered acquired knowledge to "lie outside of the realm of mental processing" (p. 2).

Whereas the manual (Kaufman & Kaufman, 2004b) suggests that the CHC model is preferred for most clinical situations, use of the Luria model is encouraged when users suspect that the administration of measures of acquired knowledge would compromise the validity of the full scale composite score and/or, if examiners in general have a firm commitment to the Luria processing tradition (p. 5).

On the KABC-II, the Luria model for school age (ages 7-18) contains eight core subtests that combine to yield a higher-order composite score called the Mental Processing Index (MPI), as well as four first-order composite scores: Sequential Processing (Short-Term Memory [Gsm]), Simultaneous Processing (Visual Processing [Gv]), Learning (Long-Term Storage and Retrieval [Glr]), and Planning (Fluid Reasoning [Gf]). The KABC-II also provides users with 10 additional supplemental subtests however, these measures do not contribute to the measurement of the MPI or the four primary composite scores and they cannot be substituted in place of the core subtests. In terms of clinical interpretation, examiners are encouraged to interpret scores in a stepwise fashion beginning with the MPI and then proceeding to the profile of first-order composite scores. However, within KABC-II interpretive resources (e.g., Kaufman & Kaufman, 2004b; Kaufman, Lichtenberger, Fletcher-Janzen, & Kaufman, 2005; Singer, Lichtenberger, Kaufman, Kaufman, & Kaufman, 2012), users are encouraged to focus most of their interpretive weight at the first-order (factor-based score) level of measurement.

Although it is suggested that the Luria and CHC interpretive frameworks are interchangeable, the test authors do not provide an explanation as to how KABC-II subtests can measure two distinct latent constructs simultaneously (Braden & Ouzts, 2005; Gallagher & Sullivan, 2011). Further complicating the matter, is the absence of information in the manual to support the structural validity of the alternative Luria model. To wit, confirmatory factor analytic (CFA) support for a five-factor CHC-based three-stratum hierarchical structure at ages 7-18 was reported in the KABC-II manual (Kaufman & Kaufman, 2004b), and Figures 8.1 and 8.2 (pp. 106-107) illustrate the standardized validation models for different configurations of the core (10 subtests) and supplemental subtests in that model. It should be noted that at ages 3-6 not all of the hypothesized CHC dimensions could be located. Specifically, the authors found it difficult to disentangle Gf (Planning) from Gv (Simultaneous Processing).

Of concern, the alternative eight subtest Luria model was never separately subjected to appropriate factor analytic procedures to uncover the latent structure measured by that configuration alone. According to Brunner, Nagy, and Wilhelm (2012), this information is important because it provides users with the statistical evidence needed to apply an interpretive framework to the standardized scores computed for a measurement instrument such as the KABC-II. This absence of this information is especially problematic as the CHC and Luria measurement models are not structurally equivalent (Cattell, 1978). It is worth noting that due to their high *g*-loadings, indicators of Gc contribute greatly to the measurement of general intelligence in contemporary ability measures (Dombrowski, Canivez, Watkins, & Beaujean, 2015). As a consequence, their omission from the Luria model may produce a weaker general factor (i.e., MPI) and the relationships between first-order dimensions may be altered (Hood, 2013). Thus, until additional empirical evidence is furnished, it cannot be assumed that the same constructs are measured well, if at all, when users elect to administer the Luria subtest configuration (Boag, 2015). Even if one were to accept the veracity of simply extrapolating the Luria structure from that of the CHC model, considerable problems would remain.

First, the test authors relied exclusively on a *constrained* CFA in which only one model, a higher-order CHC measurement model consistent with publisher theory, was fit to the KABC-II normative data at different points in the age span. That is, the fit afforded by rival models, if examined, was not reported. Although the fit statistics provided in the manual suggest that this model provided a relatively good fit to the normative data, this approach to scale validation has been criticized (Brown, 2015; Jackson, Gillaspy, & Purc-Stephenson, 2009; Keith, 2015).

According to Keith (2015), "The fact that one model fits the data reasonably well does not mean that there could not be other, *different* [emphasis added] models that fit better…The confidence with which one accepts such explanations depends, on whether other, rival explanations have been tested and found wanting" (p. 520). More concerning, the standardized path loadings between the second-order general factor and the first-order Gf (Planning) dimension for ages 7-18 were all ≥ 1.0. According to Brown (2015), these estimates suggest an impermissible solution to the data (i.e., construct redundancy). Inexplicably, this potential limitation was not disclosed in the manual.

Unfortunately, CFA investigations of the KABC-II produced in the empirical literature since its publication (e.g., Bangirana et al., 2009; Cucina & Howardson, 2016; Morgan, Rothlisberg, McIntosh, & Hunt, 2009; Reynolds, Keith, Flanagan, & Alfonso, 2013; Reynolds, Keith, Fine, Fisher, & Low, 2007) have all focused exclusively on validating the CHC structure described in the KABC-II manual using various combinations of the core and supplemental subtests associated with that particular interpretive model. In the only structural validity investigation of the Luria model that has been conducted to date, McGill and Spurgin (2015) subjected the eight core subtests to hierarchical exploratory factor analysis (EFA). Whereas, extraction criteria did not support the presence of the four factors posited by the test publisher, when a four-factor solution was forced to the KABC-II normative data, the model resulted in weak subtest loadings, a mathematically impermissible Planning factor, and theoretically inconsistent subtest migration and cross-loading between Planning and Simultaneous Processing indicators at ages 7-12 and 13-18. Additionally, it was found that most of the reliable variance in the Luria model was attributable to the higher-order MPI dimension and that the four first-order Luria dimensions likely possessed too little unique score variance for confident clinical interpretation.

Interestingly, specification of more parsimonious models (i.e., two- and three-factors) failed to produce desired simple structure and further complicated Luria model interpretation. As a result, McGill and Spurgin (2015) suggested that additional examinations using CFA techniques were needed in order to disclose the true latent structure of the KABC-II Luria model. Unfortunately, an investigation of the Luria model employing these methods has yet to be conducted suggesting that our understanding of relationships between Luria variables on the KABC-II is presently incomplete.

Purpose and Goals of the Present Study

To remediate this gap in the literature, the present study sought to test the latent factor structure of the eight subtest Luria model configuration using recommended CFA techniques with participant data from the KABC-II normative sample across the school age (ages 7-18). Although EFA and CFA are considered to be complimentary procedures, they provide answers to different empirical questions. According to Brown (2015), “CFA is more appropriate than EFA in the later stages of construct validation and test construction, when prior evidence and theory support more ‘risky’ a priori predictions regarding latent structure” (pp. 42-43).

As recommended by Keith (2015), the present analyses sought to examine the tenability of rival measurement models (e.g., higher-order model, bifactor model) that have been found to best fit the data produced from other contemporary ability measures (e.g., WISC, WJ). Previous CFA studies of the KABC-II structure have mostly examined the relationship between a higher-order general factor (*g*) and the first-order factors with the effects of *g* on the subtests fully mediated through the first-order factors (i.e., indirect hierarchical model). As an alternative, a bifactor model (Holzinger & Swineford, 1937), suggests that *g* and the group-specific factors have simultaneous direct effects on the measured variables. Although the bifactor model has been found to be a preferred solution in related research with other cognitive measures, this model has not been applied to the KABC-II. Although, the results of a recent predictive validity study (Benson, Kranzler, & Floyd, 2016) using variables from the KABC-II CHC model, suggest that it may be a preferred solution for the latent structure of the measurement instrument. As recent communications from the test publisher indicate that the KABC-II will not be revised, but will instead undergo a normative update, it is believed that the results obtained from the present investigation will be instructive for furthering our understanding of the structuring of Luria model variables and establishing evidence-based interpretive procedures for users who elect to interpret the measurement instrument from this perspective in clinical practice.

Method

Participants

Participants were children and adolescents ages 7-0 to 18-11 (*N* = 2,025) drawn from the KABC-II normative sample. Demographic characteristics are provided in detail in the KABC-II manual (Kaufman & Kaufman, 2004b). The standardization sample was obtained using stratified proportional sampling across demographic variables of age, sex, race/ethnicity, parent educational level, and geographic region. Examination of the demographic tables provided in the manual revealed a close correspondence to the 2001 U. S. census estimates across the stratification variables. The present sample was selected on the basis that it corresponded to the age ranges at which the Luria interpretive model could be fully specified.

Measurement Instrument

The KABC-II is a multidimensional test of cognitive abilities for ages 3 to 18 years. The measure is comprised of 18 subtests, eight of which contribute to the measurement of four Luria-based factor scores in the school-age battery: Sequential Processing (SQ), Simultaneous Processing (SM), Planning (P), and Learning (L). The core subtests are linearly combined to form the full scale MPI composite. All factor and composite variables on the KABC-II are expressed as standard scores with a mean of 100 and a standard deviation of 15. The total norming sample (*N* = 3,025) is nationally representative based upon 2001 U.S. census estimates. Extensive normative and psychometric data can be found in the KABC-II manual (Kaufman & Kaufman, 2004b).

Data Analyses

Confirmatory factor analysis. EQS, Version 6.2 (Bentler & Wu, 2012) was used to conduct CFA using maximum likelihood estimation. Due to the fact that the Simultaneous Processing factor is produced from different subtest combinations at ages 7-12 and 13-18, separate CFA analyses were conducted for both age groups. Consistent with previous KABC-II structural analyses, four first-order models and three hierarchical models were specified and examined at ages 7-12 and 13-18: (Model 1) one factor; (Model 2) two oblique SQ and SM factors; (Model 3) three oblique SQ, L, and combined P/SM factors; (Model 4) four oblique SQ, SM, L, and P factors; (Model 5) a four-factor bifactor model based on Model 4, (Model 6a) a four-factor higher-order model consistent with publisher theory; and (Model 6b) an alternative four-factor higher-order model with Pattern Reasoning loading on both the SM and P factors (as per McGill & Spurgin, 2015). Because the eight subtest KABC-II Luria model configuration only has two indicators for the four resulting group factors, subtests were constrained to be equal in the bifactor model to ensure identification. Beaujean (2015) and Reise (2012) have provided detailed descriptions of the salient differences between higher-order and bifactor models and the potential advantages of the later model. Examples of oblique (correlated), higher-order, and bifactor expressions of the publisher suggested four-factor Luria measurement model are outlined in Figure 1.

Model fit. To comport with best practice (e.g., Lai & Green, 2016; Mueller & Hancock, 2008), multiple indices were examined to evaluate the adequacy of model fit. Specifically, the (a) chi-square (χ^2), (b) comparative fit index (CFI), (c) root mean square error of approximation (RMSEA), (d) standardized root mean square residual (SRMR), and (e) Akaike's information criterion (AIC). Hu and Bentler (1998, 1999) have recommended a dual criterion (CFI $\geq$ 0.95 and RMSEA $\leq$ 0.06) for evaluating CFA fit statistic values to guard against both Type I and Type II errors. Higher CFI values and lower RMSEA values indicate better model fit, and these two indices were supplemented with chi-square, SRMR, and AIC values. There are no specific criteria for information-based indices like the AIC, but smaller values may indicate better approximations of the true measurement model after accounting for model complexity (Vrieze, 2012). Meaningful differences between well-fitting models were evaluated based upon the following criteria: (a) exhibit good fit according to CFI, RMSEA, and SRMR indices; (b) demonstrate a statistically significant ($p < .05$) $\Delta \chi^2$ value (for nested models); and/or (c) display the smallest AIC value (Burnham & Anderson, 2004). It should be noted that since all of the hypothesized factors in the Luria model are *just identified* (i.e., produced from only two indicators), the fit of the bifactor solution will be equivalent with its higher-order counterpart due to the imposition of equality constrains in the bifactor model. Nevertheless, Brown suggests (2015) that it may be "substantively meaningful" (p. 292) to evaluate equivalent solutions when there is evidence to suggest that they may provide a relevant explanation for the data.

Results

Eight Subtest KABC-II Luria Model Configuration for Ages 7-12

Descriptive statistics for the Luria model subtest scores for participants ages 7-12 illustrate univariate normality with the largest skewness index of .28 and the largest kurtosis index of .32. Mardia's standardized multivariate kurtosis estimate for these data was 2.33 and well within the criterion of |5.0| suggesting multivariate normality (Byrne, 2006). As a result, use of maximum likelihood estimation for CFA in the present analysis was deemed appropriate.

Model fit statistics presented in Table 1 illustrate the increasingly better fit from one to four factors; however, fit statistics indicated that the one and two factor models were inadequate (CFI < .95 and RMSEA > .06). Although the correlated three- and four-factor models both fit the data well, the oblique four-factor model (Figure 1.1) provided the best fit to these data among the first-order models ($\Delta\chi^2 = 8.89$, $\Delta df = 3$, $p < .05$). However, because the four KABC-II latent factors were highly correlated (.45-.92), a higher-order dimension is implied, rendering the oblique model an inadequate explanation for these data (Gorsuch, 1983; Thompson, 2004, Gignac, 2016).

Because of constraining each factor's loadings to equality because of empirically under-identified latent factors (SQ, SM, P, and L), Model 5 (Figure 1.3) is mathematically equivalent to Model 6a (Figure 1.2). However, in Model 6a, the standardized loading coefficient between the first-order Planning factor and the second-order *g* factor was 1.0 suggesting the presence of a Heywood case and an improper model solution (Brown, 2015). As previous factor analytic research on the KABC-II (e.g., McGill & Spurgin, 2015; Reynolds et al., 2007) has found that Pattern Reasoning was aligned with both Planning and Simultaneous Processing, an alternative higher-order model permitting this cross-loading was also examined (Model 6b). Whereas, no bifactor or higher-order factor model was superior in terms of ΔCFI .01 and ΔRMSEA .01, the alternative higher-order model (see Figure 2) fit the data the best among the four-factor models and produced the lowest AIC value of all of the models that were explicated and examined. All factor loadings in Model 6b were positive and statistically significant. Furthermore, in this model, the path loading between Planning and *g* was .98, indicating a permissible solution. As a result, it was selected as the best explanation for the Luria model structure at ages 7-12.

Effects of *g* and the Broad Abilities. Using the standardized path coefficients from Model 6b, the residualized loadings of the subtests on the second-order general (*g*) factor were obtained using the procedures outlined by Reynolds and Keith (2013). These results are reported as an online supplement in Table 2. The loadings are considered the *indirect* effects of *g* mediated through the first-order dimensions. Several of these loadings were greater than .60. Whereas, Planning and Simultaneous Processing measures were generally the best indicators of *g*, Number Recall (.401) had the lowest loading on the second-order general factor. Relatedly, Figure 3 displays how much Luria model subtest variance was explained by *g* versus how much subtest variance was explained by the first-order factors. This decomposition of variance is similar to the Schmid-Leiman (1957) procedure for EFA. Whereas, on average, *g* accounted for 31% of the variance in Luria subtests, the first-order factors explained on average 18% of the variance in the subtests. As would be expected, there was very little variance explained in the Planning measures (Story Completion, Pattern Reasoning) that was not explained by *g*.

Eight Subtest KABC-II Luria Model Configuration for Ages 13-18

Descriptive statistics for the Luria model subtest scores for participants ages 13-18 illustrate univariate normality with the largest skewness index of -.31 and the largest kurtosis index of -.23. Mardia's standardized multivariate kurtosis estimate for these data was 1.53 and well within the criterion of |5.0| suggesting multivariate normality (Byrne, 2006). As a result, use of maximum likelihood estimation for CFA in the present analysis was deemed appropriate.

Model fit statistics presented in Table 1 illustrate the increasingly better fit from one to four factors; however, fit statistics indicated that the one and two factor models were inadequate (CFI < .95 and RMSEA > .06). Although the correlated three- and four-factor models both fit the data well, the oblique four-factor model (Figure 1.1) provided the best fit to these data among the first-order models ($\Delta\chi^2 = 12.49$, $\Delta df = 3$, $p < .05$). However, because the four KABC-II latent factors were highly correlated (.51-.90), a higher-order dimension is implied, rendering the oblique model an inadequate explanation for these data (Gorsuch, 1983; Thompson, 2004, Gignac, 2016).

Because of constraining each factor's loadings to equality because of empirically under-identified latent factors (SQ, SM, P, and L), Model 5 (Figure 1.3) is mathematically equivalent to Model 6a (Figure 1.2). In contrast to the results at ages 7-12, the standardized path between Planning and *g* was .98 in Model 6a at ages 13-18, indicating a permissible solution. Nevertheless, the alternative higher-model (Model 6b) with Pattern Reasoning freed to cross-load on both Planning and Simultaneous Processing was a statistically better fit to these data than the bifactor model and the higher-order model suggested by the test publisher ($\Delta\chi^2 = 8.20$, $\Delta df = 1$, $p < .05$). Additionally, the alterative higher-order model (see Figure 4) fit the data the best and produced the lowest AIC value of all of the models that were explicated and examined. All factor loadings in Model 6b were positive and statistically significant. As a result, it was selected as the best explanation for the Luria model structure at ages 13-18.

Effects of *g* and the Broad Abilities. Using the standardized path coefficients from Model 6b, the residualized loadings of the subtests on the second-order general (*g*) factor were obtained using the procedures outlined by Reynolds and Keith (2013). These results are reported as an online supplement in Table 3. Whereas, Planning and Simultaneous Processing measures were generally the best indicators of *g*, Number Recall (.411) had the lowest loading on the second-order general factor. Relatedly, Figure 5 displays how much Luria model subtest variance was explained by *g* versus how much subtest variance was explained by the first-order factors. Whereas, on average, *g* accounted for 33% of the variance in Luria subtests, the first-order factors explained on average 19% of the variance in the subtests.

Discussion

According to Schneider (2013), CFA studies are useful for informing the clinical interpretation of ability measures in school psychology. The absence of this information for the alternative Luria interpretive model in the KABC-II manual (Kaufman & Kaufman, 2004b) and related technical resources (e.g., Kaufman & Kaufman, 2004b; Kaufman, Lichtenberger, Fletcher-Janzen, & Kaufman, 2005; Singer, Lichtenberger, Kaufman, Kaufman, & Kaufman, 2012) suggests that additional research is needed to support use of this scheme for individual decision-making in applied practice (Lilienfeld, Ammirati, & David, 2012).

Whereas a series of recent exploratory investigations failed to support the structural and predictive validity of Luria dimensions on the KABC-II (e.g., McGill & Spurgin, 2015, 2016), these studies were not dispositive for determining what psychological dimensions are measured by the Luria model. Accordingly, the purpose of the present study was to examine the structural validity of the Luria interpretive model across the school age (ages 7-18) using rival CFA procedures in an attempt to better disclose the latent structure of the Luria subtest configuration. It is believed that the present results are instructive for clinical interpretation of the KABC-II given that the structure and interpretive features for the measurement instrument are unlikely to change as a result of the forthcoming normative update by the test publisher.

The present results support a higher-order measurement model, consistent with publisher theory, at ages 7-12 and 13-18, with all eight subtests contributing to the measurement of four first-order dimensions (e.g., SQ, SM, P, and L) and a second-order general ability dimension (see Figures 2 and 4). Whereas a previous EFA investigation (McGill & Spurgin, 2015) failed to support the publisher suggested Luria measurement model, the present study found that the four-factor model was statistically a better fit to the normative data when compared to rival two- and three-factor solutions for both age groups. However, it should be noted that the higher-order model supported in the present study (see Figures 2 and 4) is a slight departure from the higher-order model suggested by the test publisher. Whereas, the model implied in the KABC-II manual, suggests that each subtest on the Luria model loads only on its theoretically assigned factor (see Figure 1.2), this model was found to be inferior to the alternative higher-order model with Pattern Reasoning being allowed to cross-load on both the Planning and Simultaneous Processing factors.

However, it should be noted that the presence of indicators that cross-load presents an interpretive challenge for practitioners as the Planning and Simultaneous Processing scores that are presented to clinicians on the KABC-II as capable of being interpreted do not account for the dimensional complexity in the Pattern Reasoning measure. That is, performance on Pattern Reasoning contributes to the Planning score but not to Simultaneous Processing despite the results from the present study suggesting that both dimensions help to explain performance on that indicator. In fact, at ages 13-18, more Pattern Reasoning variance was explained by Simultaneous Processing than its theoretically assigned factor. Unfortunately, it can be difficult for practitioners to account for these effects when interpreting observed scores at the level of the individual (Schneider, 2013).

Higher-Order Model Versus the Bifactor Model

Whereas previous investigations of related ability measures have supported a bifactor structure in which the general factor and the group-specific factors both have direct effects on observed indicators (e.g., Canivez, 2014; Canivez, Watkins, & Dombrowski, 2016; Gignac & Watkins, 2013; McGill & Dombrowski, 2016; Strickland, Watkins, & Caterino, 2015; Watkins & Beaujean, 2014), the present study indicated that a higher-order solution is preferred for the KABC-II Luria model configuration. However, due to model equivalence and the estimation problem encountered (i.e., Heywood case) in specifying the publisher suggested higher-order model (Figure 1.2) at ages 7-12, it may be argued that the bifactor model should have been given preference because of its advantages (e.g., Rodriquez, Reise, & Haviland, 2016; Reise, 2012). However, some researchers have questioned whether the bifactor model is a tenable structure for ability measures such as the KABC-II (e.g., Murray & Johnson, 2013; Reynolds & Keith, 2013). While adjudication of this issue is beyond the scope of the present discussion, it is important to note that a bifactor model is only tenable when restrictive assumptions have been met in CFA (Keith, 2015; Mansolf & Wolf, 2016). Most salient of these is that each observed variable can load on *one and only one* group-specific factor (Beaujean, 2015). However, the current results suggest a violation of perfect cluster structure in the underlying measurement model for the Luria interpretive scheme, rendering the bifactor model an untenable explanation for the present data. As noted by Chen et al. (2012), although the bifactor model can lead to greater conceptual clarity, it is not an optimal model in all conditions.

Isomorphism between Planning and *g*

As previously mentioned, in the publisher suggested higher-order model, a path coefficient of 1.0 was observed between *g* and Planning. It should be noted that similar isomorphism between these constructs was observed in all of the CFA models for the KABC-II CHC configurations that were reported in the manual (Kaufman & Kaufman, 2004b). This finding is not surprising as near perfect relations between the general factor and Fluid Reasoning-related dimensions have long been documented in the professional literature (e.g., Gustafsson & Balke, 1993; Reynolds, Keith, Flanagan, & Alfonso, 2013). Without question, the KABC-II Luria model appears to have a complex four-factor structure that does not satisfy simple structure with Pattern Reasoning having variance apportioned to more than one first-order factor. As a consequence, the problematic loadings encountered in the publisher suggested higher-order model reported here and in the manual could have been an artifact of the decision to set subtest cross-loading at zero (Golay et al., 2013). According to Asparouhov and Muthen (2009), misspecification of zero loadings in the presence of non-trivial cross-loading can lead to distorted structural relationships in a higher-order measurement model.

Effects of *g* Versus the Broad Abilities

Decomposed variance estimates based on the alternative higher-order model for ages 7-12 and 13-18 displayed in Figures 3 and 5 illustrate the portions of subtest variance in the Luria model that can be sourced to the second-order *g* factor as well the four first-order group factors (SQ, SM, P, and L). Whereas, greater proportions of subtest variance were associated with the indirect effects of *g*, all of the individual group factors, with the exception of Planning, explained meaningful proportions of variance in Luria indicators for both age groups. As an example, at ages 13-18, Sequential Processing (36%) accounted for more of the total variance in Number Recall and Word Order than general intelligence (24%). Additionally, unique variance estimates ranged from 33%-66% for ages 7-12 and 21%-65% at ages 13-18 indicating that a relatively large portion of performance on these measures is attributable to a combination of specificity and measurement error (Brown, 2015).

Limitations and Future Directions

This study is not without limitations that should be considered when interpreting the results. The most important limitation is the use of an archived standardization sample. Although the sample was relatively large and nationally representative, additional research to determine if these results are invariant across different samples and/or clinical settings would be instructive for informing clinical interpretation of the KABC-II.

Whereas Canivez and Kush (2013) have suggested that specification of post-hoc adjustments to measurement models in CFA are akin to *fishing expeditions*, the imposition of the additional parameter in the alternative higher-order model was not arbitrary and was informed by evidence provided by the Luria model EFA conducted by McGill and Spurgin (2015) in which Pattern Reasoning was found to be *aligned* with multiple latent factors. Additionally, it should be noted that in a previous CFA of the KABC-II CHC total battery configuration with all 18 subtests (Reynolds et al., 2007), a similar model including this same cross-loading was found to best fit the KABC-II normative data. As a result, it is believed that failing to explicate a model that included this parameter would have rendered the present study incomplete.

Finally, despite the fact that the verbiage associated with the Luria interpretive model was used to identify the latent dimensions that were located in these CFAs, it is psychometrically implausible for the Luria model subtests to measure two distinct, and theoretically divergent, constructs simultaneously (Braden & Ouzts, 2005). As a result, clinicians using the KABC-II are encouraged to employ a consistent theoretical framework (i.e., CHC) to the scores obtained from the measurement instrument, no matter which subset configuration they elect to administer (i.e., Flanagan, Alfonso, Ortiz, & Dynda, 2013).

Conclusions and Implications for Clinical Practice

Over the last decade, unrestricted EFA methods have been eclipsed by more restrictive CFA methods when examining the structural validity of psychological measures (Reynolds & Keith, 2013). As stated by Gorsuch (2003), “the ultimate arbiter in science is well established: replication” (p. 153). As EFA and CFA provide answers to different empirical questions, contradictory results are commonly reported within the cognitive assessment literature (Canivez, 2013). The current study supported a structure for the alternative Luria model subtest configuration for the KABC-II at ages 7-18 that largely cohered with publisher theory. Although the CHC model is more commonly utilized to interpret KABC-II scores in contemporary practice, these results suggest that clinicians may employ that subtest configuration with greater confidence in clinical assessments. Whereas, in general, these results are relatively consistent with the Luria model EFA conducted by McGill and Spurgin (2015), the present study provided more evidence to support clinical interpretation of the Luria model factor-based scores, and in some cases, individual subtests.

Nevertheless, the present study illustrates well that apportioned variance in Luria model subtests can be sourced to multiple psychological constructs. The meaningful portions of variance accounted for by first-order Luria dimensions indicate that additional consideration of the factor-based scores (SQ, SM, L, and P) beyond the MPI composite may be warranted as long as due consideration is given to the indirect influence of the general ability factor (*g*) in all of those indicators. However, clinicians are encouraged to interpret the Planning and Simultaneous Processing scores with caution given the fact that Pattern Reasoning was found to cross-load on both of these factors. The factor-based scores corresponding with those measures do not account for this complexity and thus there is a risk of under *or* overestimating the contributions of these dimensions in explaining performance on this indicator.

References

Asparouhov, T., & Muthen, B. (2009). Exploratory structure equation modeling. *Structural Equation Modeling, 16,* 397-438. doi: 10.1080/10705510903008204

Bangirana, P., Seggane, M., Allebeck, P., Giordani, B., Chandy, C. J., Opoka, O. R., Byarugaba, J., Ehnvall, A., & Boivin, M. J. (2009). A preliminary examination of the construct validity of the KABC-II in Ugandan children with a history of cerebral malaria. *African Health Sciences, 9,* 186-192. Retrieved from http://www.ajol.info/index.php/ahs/article/view/49010/35360

Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence, 3,* 121-136. doi: 10.3390/jintelligence3040121

Benson, N. F., Kranzler, J. H., & Floyd, R. G. (2016). Examining the integrity of measurement of cognitive abilities in the prediction of achievement: Comparisons and contrasts across variables from higher-order and bifactor models. *Journal of School Psychology, 58,* 1-19. doi: 10.1016/j.jsp.2016.06.001

Bentler, P. M., & Wu, E. J. C. (2012). *EQS for Windows*. Encino CA: Multivariate Software.

Boag, S. (2015). Personality assessment, 'construct validity', and the significance of theory. *Personality and Individual Differences, 84,* 36-44. doi: 10.1016/j.paid.2014.12.039

Braden, J. P. & Ouzts, S. M. (2005). Review of Kaufman Assessment Battery for Children, Second Edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 517-520). Lincoln: University of Nebraska.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford Press.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs.

Journal of Personality, 80, 796-846. doi: 10.1111/j.1467-6494.2011.00749.x

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33,* 261-304. doi: 10.1177/0049124104268644

Byrne, B. M. (2016). *Structural equation modeling with EQS: Basic concepts, applications, and programming* (2nd ed.). New York: Erlbaum.

Canivez, G. L. (2013). Psychometric versus actuarial interpretation of intelligence and related aptitude batteries. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.). *The* Oxford handbook of child psychological assessment (pp. 84-112). New York: Oxford University Press.

Canivez, G. L. (2014). Construct validity of the WISC-IV with a referred sample: Direct versus indirect hierarchical structures. *School Psychology Quarterly, 29,* 38-51. doi: 10.1037/spq0000032

Canivez, G. L., & Kush, J. C. (2013). WISC-IV and WAIS-IV structural validity: Alternate methods, alternate results. Commentary on Weiss et al. (2013a) and Weiss et al. (2013b). *Journal of Psychoeducational Assessment, 31,* 157-169. doi: 10.1177/0734282913478036

Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2016). Structural validity of the Wechsler Intelligence Scale for Children-Fifth Edition: Confirmatory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*. Advance online publication. doi: 10.1037/pas0000358

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J. –P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality, 80,* 219–251. doi:10.1111/j.1467–6494.2011.00739.x

Dombrowski, S. C., Canivez, G. L., Watkins, M. W., & Beaujean, A. (2015). Exploratory bifactor analysis of the Wechsler Intelligence Scale for Children-Fifth Edition with the 16 primary and secondary subtests. *Intelligence, 53,* 194-201. doi: 10.1016/j.intell.2015.10.009

Flanagan, D. P., Alfonso, V. C., Ortiz, S. O., & Dynda, A. M. (2013). Cognitive assessment: Progress in psychometric theories of intelligence, the structure of cognitive ability tests, and interpretive approaches to cognitive test performance. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.). The Oxford handbook of child psychological assessment (pp. 239-285). New York: Oxford University Press.

Gallagher, S. L., & Sullivan, A. L. (2011). The Kaufman Assessment Battery for Children, Second Edition. In A. S. Davis (Ed.), *Handbook of pediatric neuropsychology* (pp. 343-352). New York: Springer.

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bi-factor model tends to fit better. *Intelligence, 55,* 57-68. doi: 10.1016/j.intell.2016.01.006

Gignac, G. E., & Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research, 48,* 639-662. doi: 10.1080/00273171.2013.804398

Golay, P., Reverte, I., Rossier, J., Favez, N., & Lecerf, T. (2013). Further insights on the French

WISC-IV factor structure through Bayesian structural equation modeling. *Psychological Assessment, 25,* 496-508. doi: 10.1037/a0030676

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: John Wiley.

Gustafsson, J. –E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research, 28,* 407-434.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika, 2,* 41-54. doi:10.1007/BF02287965

Hood, S. B. (2013). Psychological measurement and methodological realism. *Erkenntis, 78,* 739-761. doi: 10.1007/s10670-013-9502-z

Hu, L.T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to under parameterized model misspecification. *Psychological Methods, 3*, 424–453. doi:10.1037/1082- 989X.3.4.424

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55. doi:10.1080/ 10705519909540118

Jackson, D. L., Gillapsy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14,* 6-23. doi: 10.1037/a0014694

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004a). *Kaufman Assessment Battery for Children-Second Edition*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004b). *Kaufman Assessment Battery for Children-Second Edition manual*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., & Kaufman, N. L. (2005). *Essentials of KABC-II assessment*. Hoboken, NJ: John Wiley.

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive *g* and academic achievement *g* one and the same *g*? An exploration on the Woodcock-Johnson and Kaufman tests. *Intelligence, 40,* 123-138. doi: 10.1016/j.intell.2012.01.009

Keith, T. Z. (2015). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling* (2nd ed.). New York: Routledge.

Lai, K., & Green, S. B. (2016). The problem with having two watches: Assessment of fit when RMSEA and CFI disagree. *Multivariate Behavioral Research, 51,* 220-239. doi: 10.1080/00273171.2015.1134306

Lilienfeld, S. O., Ammirati, R., & David, M. (2012). Distinguishing between science pseudoscience in school psychology: Science and scientific thinking as safeguards against human error. *Journal of School Psychology, 50,* 7-36. doi: 10.1016/j.jsp.2011.09.006

Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper Row.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. London: Penguin Books.

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthgonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research, 51,* 698-717. doi: 0.1080/00273171.2016.1215898

McGill, R. J., & Dombrowski, S. C. (2016). What does the WRAML2 core battery measure? Utilizing exploratory and confirmatory techniques to disclose higher-order structure. *Assessment*. Advance online publication. doi: 10.1177/1073191116677799

McGill, R. J., & Spurgin, A. R. (2015). Exploratory higher order analysis of the Luria interpretive model on the Kaufman Assessment Battery for Children-Second Edition (KABC-II) school-age battery. *Assessment*. Advance online publication. doi: 10.1177/1073191115614081

McGill, R. J., & Spurgin, A. R. (2016). Assessing the incremental value of KABC-II Luria model scores in predicting achievement: What do they tell us beyond the MPI? *Psychology in the Schools, 53,* 677-689. doi: 10.1002/pits.21940

Morgan, K. E., Rothlisberg, B. A., McIntosh, D. E., & Hunt, M. S. (2009). Confirmatory factor analysis of the KABC-II in preschool children. *Psychology in the Schools, 46,* 515-525. doi: 10.1002/pits.20394

Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. W. Osborne (Ed.), *Best practices in quantitative methods*. Thousand Oaks, CA: Sage.

Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence, 37,* 487-494. doi: 10.1016/j.intell.2013.06.004

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47,* 667-696. doi:10.1080/00273171.2012.715555

Reynolds, M. R., & T. Z. Keith (2013). Measurement and statistical issues in child assessment research. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.). *The Oxford handbook of child psychological assessment* (pp. 48-83). New York: Oxford University Press.

Reynolds, M. R., Keith, T. Z., Fine, J. G., Fisher, M. E., & Low, J. (2007). Confirmatory factor structure of the Kaufman Assessment Battery for Children-Second Edition: Consistency with Cattell-Horn-Carroll theory. *School Psychology Quarterly, 22,* 511-539. doi: 10.1037/1045-3830.22.4.511

Reynolds, M. R., Keith, T. Z., Flanagan, D. P., & Alfonso, V. C. (2013). A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. *Journal of School Psychology, 51,* 535-555. doi: 10.1016/j.jsp.2013.02.003

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98,* 223-237. doi: 10.1080/00223891.2015.1089249

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22,* 53-61. doi: 10.1007/BF02289209

Schneider, W. J. (2013). What if we took our models seriously? Estimating latent scores in individuals. *Journal of Psychoeducational Assessment, 31,* 186-201. doi: 10.1177/0734282913478046

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99-144). New York: Guilford Press.

Singer, J. K., Lichtenberger, E. O., Kaufman, J. C., Kaufman, A. S., & Kaufman, N. L. (2012). The Kaufman Assessment Battery for Children-Second Edition and the Kaufman Test of

Educational Achievement-Second Edition. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 269-296). New York: Guilford Press.

Sotelo-Dynega, M., & Dixon, S. G. (2014). Cognitive assessment practices: A survey of school psychologists. *Psychology in the Schools, 51,* 1031-1045. doi: 10.1002/pits.21800

Strickland, T., Watkins, M. W., & Caterino, L. C. (2015). Structure of the Woodcock-Johnson III cognitive tests in a referral sample of elementary school children. *Psychological Assessment, 27,* 689-697. doi: 10.1037/pas0000052

Styck, K. M., & Watkins, M. W. (2016). Structural validity of the WISC-IV for students with learning disabilities. *Journal of Learning Disabilities, 49,* 216-224. doi: 10.1177/0022219414539565

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

Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods, 17,* 228 -243. doi: 10.1037/a0027127

Watkins, M. W., & Beaujean, A. A. (2014). Bifactor structure of the Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition. *School Psychology Quarterly, 29,* 52-63. doi: 10.1037/spq0000038

Table 1

Confirmatory Factor Analysis Fit Statistics for KABC-II Eight Subtest Luria Configuration for Normative Sample Participants Ages 7-18 (N = 2,025)

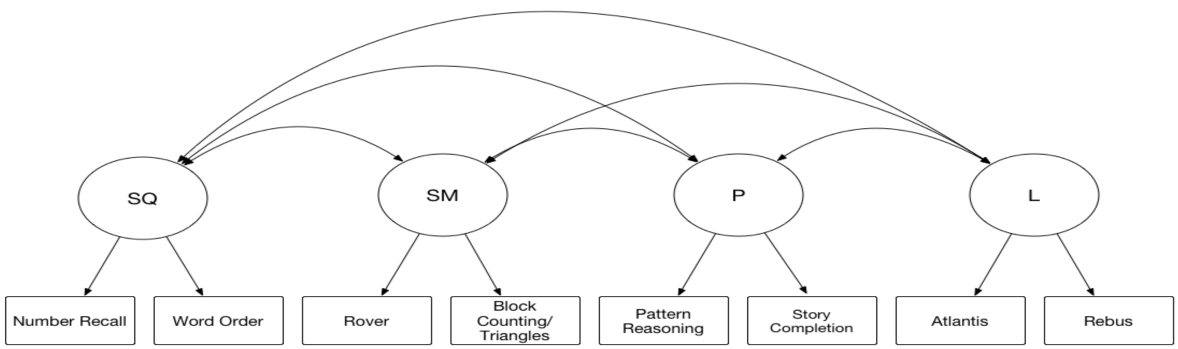
Model	χ^2	*df*	*p*	CFI	SRMR	RMSEA	90% CI RMSEA	AIC
Ages 7-12 (*n* = 1,142)								
1. One factor (*g*)	383.98	20	.00	.840	.067	.126	[.115, .137]	343.98
2. Two oblique factors (SQ, SM)	319.04	19	.00	.868	.061	.118	[.106, .129]	281.04
3. Three oblique factors (SQ, L, SM/P)	35.67	17	.01	.992	.018	.031	[.016, .045]	-1.67
4. Four oblique factors (Luria)	26.78*	14	.02	.994	.015	.028	[.011, .044]	-1.23
5. Bifactor[a] version of Model 4	27.69	16	.03	.995	.016	.025	[.007, .041]	-4.31
6a. Higher-order	27.69	16	.03	.995	.016	.025	[.007, .041]	-4.31
6b. Higher-order (PR on SM & P)	**25.15**	**15**	**.04**	**.996**	**.015**	**.024**	**[.002, .040]**	**-4.85**
Ages 13-18 (*n* = 883)								
1. One factor (*g*)	324.47	20	.00	.846	.067	.131	[.119, .144]	284.47
2. Two oblique factors (SQ, SM)	269.36	19	.00	.873	.059	.122	[.109, .135]	231.36
3. Three oblique factors (SQ, L, SM/P)	59.82	17	.00	.978	.027	.053	[.039, .068]	25.82
4. Four oblique factors (Luria)	47.33*	14	.00	.983	.024	.052	[.036, .069]	19.33
5. Bifactor[a] version of Model 4	53.86	16	.00	.981	.027	.052	[.037, .067]	21.86
6a. Higher-Order	53.86	16	.00	.981	.027	.052	[.037, .067]	21.86
6b. Higher-order (PR on SM & P)	**45.66****	**15**	**.00**	**.984**	**.025**	**.048**	**[.033, .064]**	**15.67**

Note. KABC-II = Kaufman Assessment Battery for Children-Second Edition. CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; AIC = Akaike information criterion. *g* = general intelligence, SQ = sequential processing, L = learning, SM = simultaneous processing, P = planning, PR = pattern reasoning. In the Luria oblique four-factor model for ages 7-12, correlations ranged from .45 to .92. In the Luria oblique four-factor model for ages 13-18, correlations ranged from .51 to .90.

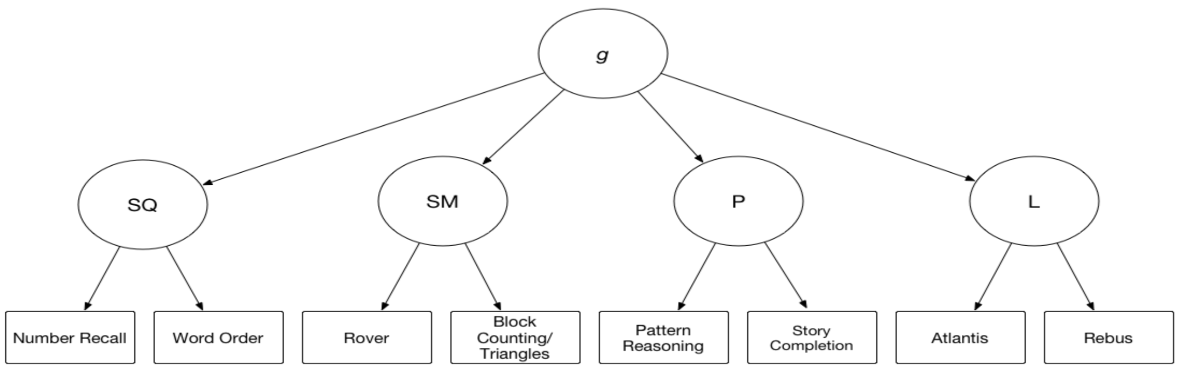
[a] Group-specific factors with less than three indicators were constrained in order to ensure identification.

[*] Statistically different ($p < .05$) from previous models.

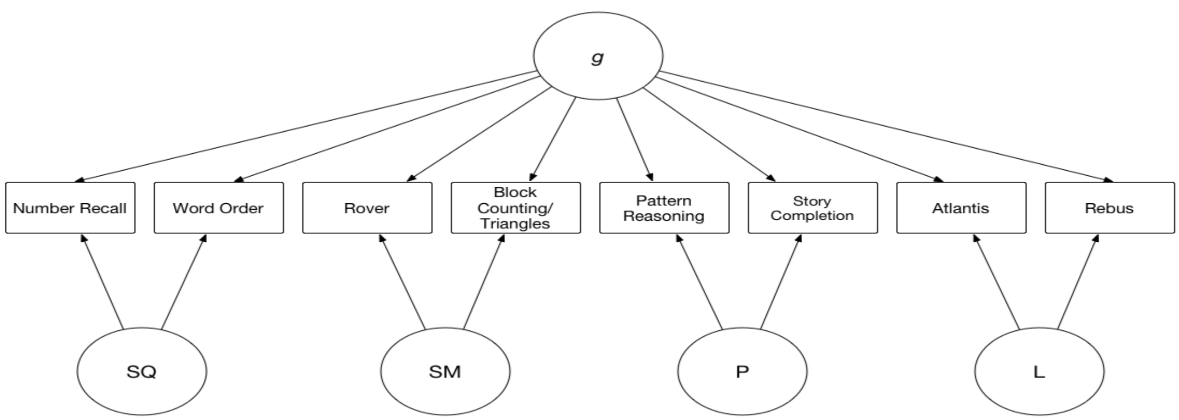
[**] Statistically different ($p < .05$) from previous two models.



1.1



1.2



1.3

Figure 1. Examples of three different multidimensional measurement models for the eight subtest Luria interpretive model for the KABC-II ages 7-18 with four latent first-order Factors. Model 1.1 is an oblique (correlated factors model, Model 1.2 is the higher-order model with one second-order factor (*g*) influencing the observed variables indirectly through the first-order factors, and Model 1.3 is the bifactor model with one general factor and four groups-specific factors all with direct effects on the observed variables. SQ = Sequential Processing, SM = Simultaneous Processing, P = Planning, L = Learning.

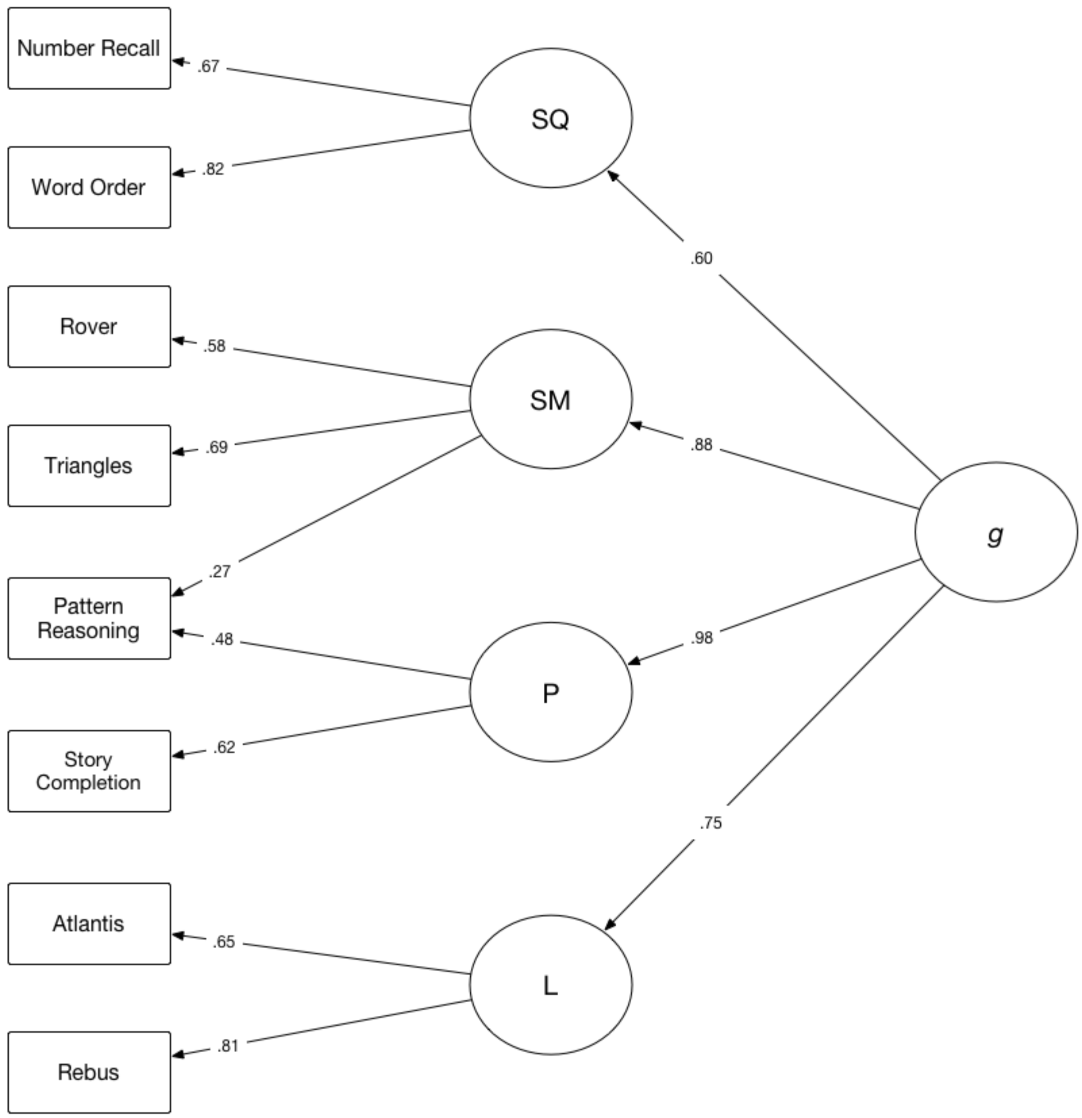


Figure 2. Alternative higher-order measurement model (6b) with standardized loading coefficients for the KABC-II Luria eight subtest configuration for ages 7-12. *g* = general intelligence; SQ = Sequential Processing; SM = Simultaneous Processing; P = Planning; L = Learning. For clarity, error terms are omitted.

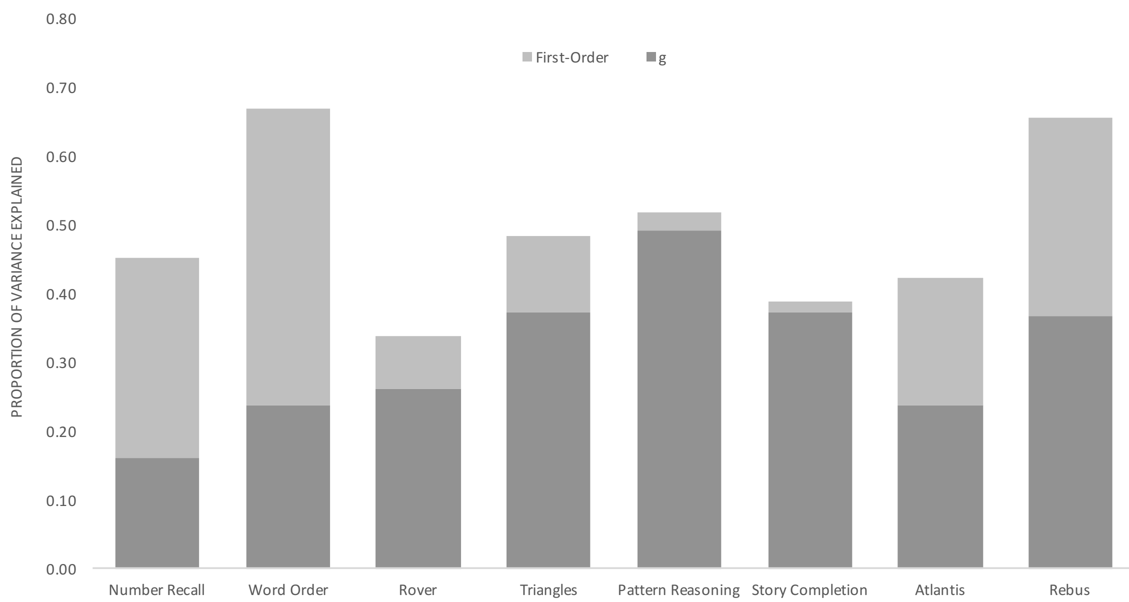


Figure 3. Sources of Variance for the Eight Subtest KABC-II Luria Model Battery Ages 7-12 (*N* = 1,142) According to a four-factor higher-order model. *g* = general intelligence.

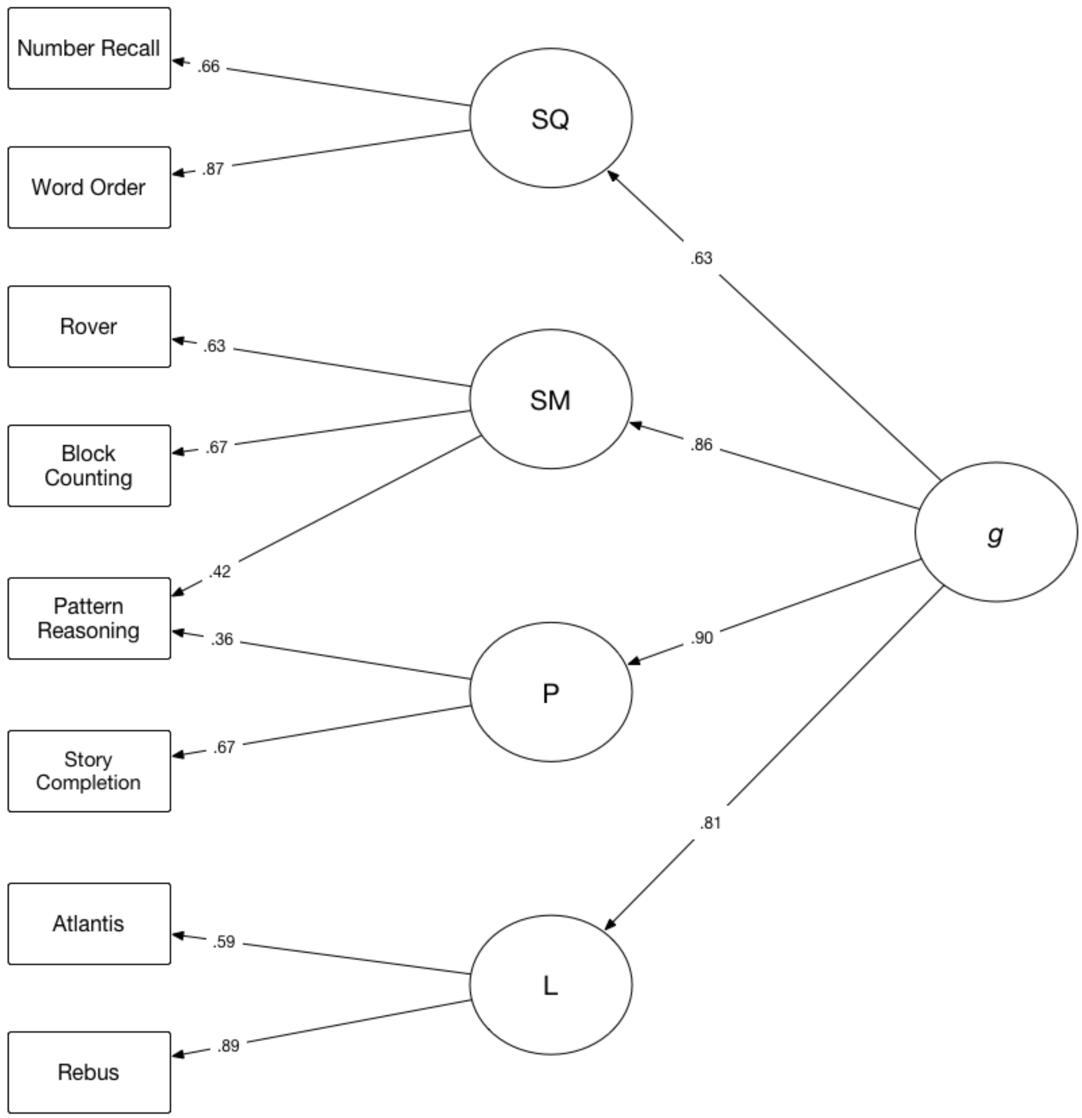


Figure 4. Alternative higher-order measurement model (6b) with standardized loading coefficients for the KABC-II Luria eight subtest configuration for ages 13-18. *g* = general intelligence; SQ = Sequential Processing; SM = Simultaneous Processing; P = Planning; L = Learning. For clarity, error terms are omitted.

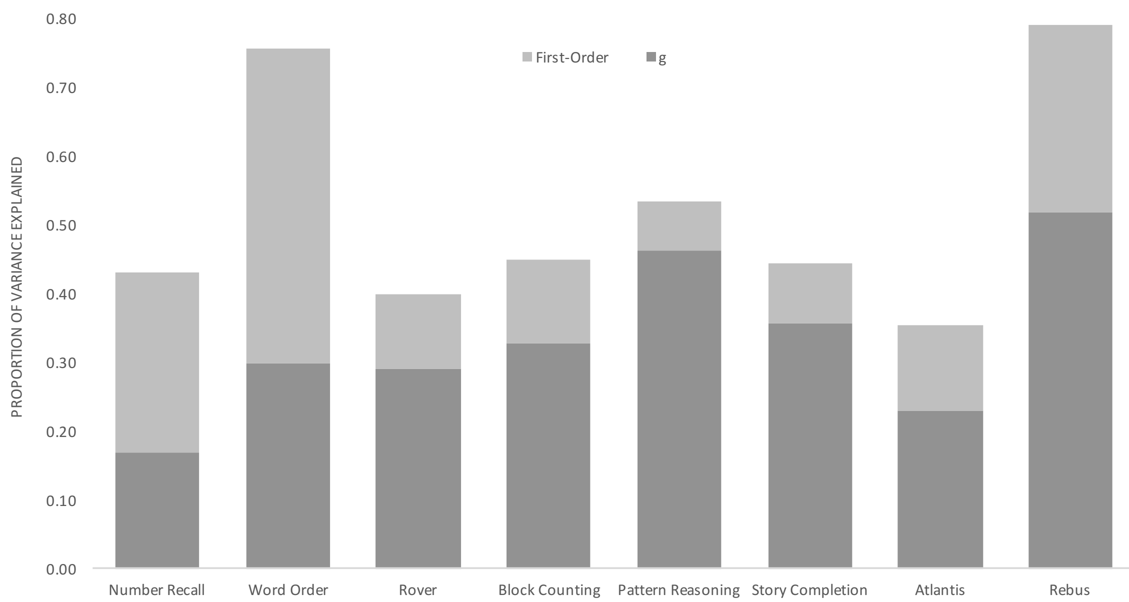


Figure 5. Sources of Variance for the Eight Subtest KABC-II Luria Model Battery Ages 13-18 (N = 883) According to a four-factor higher-order model. g = general intelligence.