Published in final edited form as:
Dyslexia. 2022 February ; 28(1): 20–39. doi:10.1002/dys.1699.

Phonological Working Memory and Central Executive Function Differ in Children with Typical Development and Dyslexia

Mary Alt,
Department of Speech, Language, and Hearing Sciences, University of Arizona, Tucson, AZ

Annie Fox,
Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, MA

Roy Levy,
T. Denny Sanford School of Social and Family Dynamics, Arizona State University, Tempe, AZ

Tiffany P. Hogan,
Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, MA

Nelson Cowan,
Department of Psychological Sciences, University of Missouri-Columbia, Columbia, MO

Shelley Gray
Department of Speech and Hearing Science, Arizona State University, Tempe, AZ

Abstract

The primary purpose of this study was to compare the working memory performance of monolingual English-speaking 2nd−grade children with dyslexia (N=82) to 2nd grade children with typical development (N = 167). Prior to making group comparisons, it is important to demonstrate invariance between working memory models in both groups or between-group comparisons would not be valid. Thus, we completed invariance testing using a model of working memory that had been validated for children with typical development (Gray et al., 2017) to see if it was valid for children with dyslexia. We tested three types of invariance: configural (does the model test the same constructs?), metric (are the factor loadings equivalent?), and scalar (are the item intercepts the same?). Group comparisons favored the children with typical development across all three working memory factors. However, differences in the Focus-of-Attention/Visuospatial factor could be explained by group differences in non-verbal intelligence and language skills. In contrast, differences in the Phonological and Central Executive working memory factors remained, even after accounting for non-verbal intelligence and language. Results highlight the need for researchers and educators to attend not only to the phonological aspects of working memory in children with dyslexia, but also to central executive function.

Corresponding Author: Mary Alt, 1131 E. 2nd St., Tucson, AZ, 85721, 520-626-6180, fax: 520-626-1364, malt@arizona.edu.

Conflict of Interest: No authors have any real or potential conflicts of interest

Working memory can be defined as an individual’s capacity for retaining a limited amount of information in an accessible form that allows it to be used or manipulated to carry out various cognitive tasks (Cowan, 2017). It underpins important academic functions (e.g., Cowan, 2014; Knoop-van Campen et al., 2018; Peng et al., 2016; 2018; Rhodes et al., 2016) with research suggesting that aspects of working memory are impaired in at least a subset of children with dyslexia (e.g., Gray et al., 2019; Jeffries & Everatt, 2004; Swanson & Sachse-Lee, 2001). Prior to making between-group comparisons that would suggest working memory impairment, however, we require evidence that the same working memory construct is being measured across groups. Invariance analyses test this very thing: they mathematically determine if a model fits similarly for two groups. If models are not invariant across groups, results of between-group comparisons could be invalid, uninterpretable (e.g., Fisher & Karl, 2019), and include systematic measurement error (e.g., Yoon & Lai, 2018). Therefore, the first purpose of this study was to determine whether a working memory model that fit the data for second-grade children with typical development (TD; Gray et al., 2017) would be invariant when we tested the same model with a group of second-grade children with dyslexia who completed the same working memory tasks. We used a multi-step process to test configural invariance (equivalence of model form), metric invariance (equivalence of factor loadings) and scalar invariance (equivalence of item intercepts). With invariance established, we proceeded to compare factor means on working memory measures for the groups with TD and dyslexia and to determine whether non-verbal intelligence and omnibus oral language measures could account for potential differences in the factor means (see Figure 1).

Keywords

dyslexia; working memory; children; invariance; model; cognition

Academic Outcomes: The Influence of Working Memory

Working memory could directly influence academic achievement in reading (e.g., remembering the words at the beginning of the sentence, as you process the end of the sentence to construct meaning) and math (e.g., holding in mind which digit goes in the ones place and which is carried to the tens place). Indeed, Maehler and Schuchardt (2016b) found that working memory was a stronger predictor of academic outcomes than IQ for 98 third-grade children with and without learning disabilities. Children who had low academic achievement also had low working memory performance, regardless of their IQ. Conversely, children who had adequate academic achievement also had adequate working memory performance, regardless of their IQ. Thus, evidence suggests that working memory uniquely impacts school achievement.

Brain Differences in Dyslexia

While there are no suggestions that working memory might function in a fundamentally different way in children with dyslexia compared to their peers with TD, it is possible. Specifically, there are neuroimaging data that point to both structural and functional differences in the brains of children and adults with dyslexia on reading tasks compared to peers with no impairment (e.g., D’Mello & Gabrieli, 2018). Given the link between reading and working memory (e.g., Knoop-van Campen et al., 2018; Peng et al., 2018;

comprehension specifically: Cain et al., 2004; Daneman & Carpenter, 1980), it is possible that neurological differences in children with dyslexia could lead to a different construct of working memory. Some of the differences in people with and without dyslexia that D'Mello and Gabrieli (2018) highlighted in their review were an under activation of the left occipito-temporal cortex and the left tempoparietal cortex, in contrast to heightened activation in parts of the left interior frontal cortex. They noted that the same areas showing functional differences also showed structural differences, including reduced gray matter and decreased white matter connections.

Working Memory Models in Children

Several models of working memory have been tested in children with agreement regarding model components growing closer over time (Gray et al., 2017; Hornung, 2011). Baddeley and Hitch's (1974) three-component model included a Central Executive (the component that directs attention), a Phonological Loop (a place for rehearsal and storage of phonological and auditory input), and a Visuospatial Sketchpad (a place for rehearsal and storage of visual and spatial input). Cowan's embedded processes model (1995a, 1999, 2001) included a Central Executive, a Phonological Storage and Rehearsal component, and a Focus of Attention component (which is capacity limited and used for visual information). Baddeley (2000) proposed a four-component model that included an Episodic Buffer, which was a component that allowed for binding of information (e.g., a word label to a word referent). While researchers have tested models of working memory in children (e.g., Alloway et al., 2004; 2006; Bayliss et al., 2003; Campos et al., 2013; Hornung, 2011; Gathercole et al., 2004; Giofré et al., 2013, Michalczyk et al., 2013; Nadler & Archibald, 2014), few of these groups have included measures appropriate to test the episodic buffer, or were missing other measures that did not allow them to test all potential working memory factors (e.g., Central Executive, Visuospatial Sketchpad). Gray et al. (2017) used a comprehensive battery that allowed them to fully test models of working memory. They used data from 168 typically-developing second-graders and found that a model that combined Cowan's embedded process model and Baddeley and Hitch's three-factor model fit well. The resulting three factors were: (a) Central Executive, (b) Focus of Attention/Visuospatial Sketchpad (Focus/Visuospatial), and (c) Phonological Storage and Rehearsal/Phonological Loop (Phonological). The model testing Baddeley's (2000) four-component model that included an Episodic Buffer failed to converge (although it can be argued that future research using not only cross-modal and cross-feature binding but also semantic elements to represent the episodic buffer could produce a more positive result; see Baddeley et al., 2021). Thus, for children with TD, there was evidence for a combined Cowan/Baddeley and Hitch model of working memory.

Working Memory and Dyslexia

Despite the lack of invariance testing, numerous studies have identified working memory deficits in children with dyslexia, with all components of working memory at risk, regardless of modality. There are numerous reports of poor verbal or phonological working memory (e.g., Alloway et al., 2017; Berninger et al., 2006; 2008), reports of poor visual working memory (e.g., Menghini et al., 2011), and difficulty binding visual and phonological information (e.g., Litt et al, 2019; Toffalini et al., 2018). Findings relative to deficits in the central executive are more tenuous with examples of individual differences (e.g., Berninger et al., 2008) or deficits that disappear once other factors are controlled (e.g., Schuchardt et al., 2008). Thus, we need to tread carefully when generalizing about working memory abilities in people with dyslexia. First, having well-defined subject groups is important because different types of working memory deficits may be attributed to dyslexia versus attention deficit hyperactivity disorder (Maehler & Schuchardt, 2016a), specific language impairment (Wong et al., 2017), and dyscalculia (Schuchardt et al., 2008). If subject groups are not carefully defined to rule out comorbid conditions, there is the potential that poor performance may be attributable to a different cause (e.g., attention issues) instead of dyslexia. Second, not all children with dyslexia have poor working memory skills. Gray et al. (2019) examined the working memory profiles of hundreds of second graders, including those with dyslexia. Four working memory profiles emerged. Importantly, children with dyslexia were represented in each profile, including those with no working memory concerns. While 20% of children with dyslexia were in the low-profile category, 24% were in the high-profile category. This distribution is important in that, for some children with dyslexia, working memory skills might be used to help compensate for, or reduce the impact of, language or reading deficits.

Current Study

The first purpose of this study was to establish model invariance (Putnick & Bornstein, 2016) across a group of second graders with TD (Gray et al., 2017) and a group of second graders with dyslexia who completed the same working memory battery. If we find similar fit for both groups when the model structure, factor loadings, and intercepts are tested, this suggests between-group comparisons for working memory factors can be interpreted appropriately. Our hypothesis was that the structure of the model (configural invariance) would be invariant. Although we hoped for metric and scalar invariance so we could make comparisons, it was unclear what to expect for the factor loadings (metric invariance) and intercepts (scalar invariance) based on the literature.

Our second purpose, given invariant findings, was to compare between-group factor means to determine whether the group with dyslexia scored significantly lower than the group with TD in any working memory factor. We also examined the patterns of correlations among working memory factors for each group and whether potential between-group differences could be accounted for by language (i.e., omnibus oral language) or nonverbal IQ measures.

Method

Participants

After receiving Institutional Review Board approval at all sites (metropolitan sites in Arizona (2), Massachusetts (1) and Nebraska (1)), we enrolled 249 2nd graders with typical oral language skills and nonverbal intelligence. Participants were part of a larger study of working memory and word learning that also included children with developmental language disorder (DLD), children with dyslexia and DLD, and bilingual Spanish-English speaking children with TD. The 249 children included in this study were 167 children with typical development (TD) (ages 6;8 – 9;2 years; months) and 82 children with dyslexia (ages

7;1 – 8;11 years; months). The TD group was the same as reported in Gray et al., 2017. In total, there were 141 girls and 108 boys, and all were monolingual English-speakers. Thirty-six children were Hispanic, 209 were non-Hispanic, and four were of unknown ethnicity. Based on parent report, 2.4% of children were Native American, 1.2 % were Asian, 2.4% were Black, 78.71% were White, 12.85% reported more than one race, and 2.4% did not report a race.

Children in both groups had to meet the following inclusionary criteria: (1) enrolled in second grade; (2) had not repeated a grade, (3) passed a bilateral hearing screening, near vision acuity screening, and color vision screening; (4) no history of neurodevelopmental disorders (e.g., ADHD, autism spectrum disorder) per parental report; (5) monolingual English speaker; (6) achieved a standard score above 74 on the Nonverbal Index of the Kaufman Assessment Battery for Children - Second Edition (K-ABC II; Kaufman & Kaufman, 2004) to rule out intellectual disability; (7) achieved a standard score above 87 on the core language composite of the Clinical Evaluation of Language Fundamentals – Fourth Edition (CELF-4; Semel et al., 2003) to rule out developmental language disorder, and (8) scored above the 30th percentile on the Goldman-Fristoe Test of Articulation – Second Edition (GFTA-2; Goldman & Fristoe, 2000). Children who scored below the 30th percentile were included in the study if their articulation errors did not affect their performance on experimental production measures (n =25).

Additionally, children in the TD group were required to have no history of special education services and to achieve a grade level composite standard score above 95 on the Test of Word Reading Efficiency – Second Edition (TOWRE-2; Torgesen et al., 2012).

Children with dyslexia were required to receive a grade level composite standard score of 88 or below on the TOWRE-2. Some of these children were receiving special education services for reading, but this was not a requirement for inclusion given that many children with dyslexia are not diagnosed in early elementary school. While dyslexia has genetic and neurological underpinnings (e.g., Centanni, 2020), it is diagnosed behaviorally. The TOWRE-2 assesses two hallmarks of dyslexia: poor word reading fluency and poor decoding. Our cut-point (20th percentile) for including children in the group with dyslexia is within the midrange of suggested percentile cut-points in the literature (e.g., Alt et al., 2017), and our use of a score buffer (i.e., excluding children with standard scores between 89 and 94) as well as our use of oral language, nonverbal IQ, and articulation measures, helped to ensure that we were not including children who likely did not have dyslexia, but who demonstrated poor reading in conjunction with other problems such as low oral language or listening comprehension. Table 1 includes details for the inclusionary measures and descriptive data.

Procedures

All sessions were conducted individually in a quiet location, such as the child's school, home, a laboratory, or a local library. Children also participated in a series of computer-based experimental tasks from the Comprehensive Assessment Battery for Children – Working Memory (CABC-WM; Cabbage et al., 2017) and the Comprehensive Assessment Battery for Children – Word Learning (CABC- WL; Gray et al., n.d.; word learning data not included in these analyses). Each child participated in approximately 7 sessions that lasted 1–2 hours each.

Experimental Tasks

The CABC-WM (Cabbage et al., 2017) consists of 13 working memory tasks, which were administered in a random order across and within research sessions. Children were seated in front of a touchscreen computer that presented pirate-themed games. They were instructed to help their chosen pirate avatar solve problems. Each task began with instructions and a demonstration by a pirate guide, followed by a series of training trails that children were required to pass. If children did not pass the training, the pirate guided them to the next game, and an imputed score was used for that game. Once children who passed training began a game, they were exposed to all the stimuli for that game; there were no accuracy cut-offs that terminated a game early. Children did not receive explicit feedback as to whether their responses were correct or incorrect, but were rewarded with virtual coins that they could spend in a virtual pirate store. We selected the CABC-WM because it was designed to assess multiple components of working memory. We designed it for children in a way that minimized the linguistic load of each task (e.g., using difficult-to-name polygons v. squares for visual tasks) to reduce the possibility of confounding working memory performance with linguistic performance. Details can be found in Cabbage et al. (2017), but the basics of each task are outlined below.

Central Executive Tasks—There were three Central Executive tasks: Number Updating, N-Back Auditory, and N-Back Visual. In Number Updating, children had to keep track of how many yo-yos and teddy bears were being ordered from a toy factory so that the babies who were expecting the toys would not cry. Children saw two black squares that contained a certain number of yo-yos or teddy bears and the digits that corresponded to the number of items in each square. Then, they would see two new red squares, one of which contained “+1”, the other of which was empty. Then, children would see two empty squares outlined in green, which cued them to report on the updated number of yo-yos and teddy bears. Children continually updated the number of items for five trials in a row, with a total of three blocks of five trials. The primary scoring metric was a lenient one, so that if a child made an error on one trial, but subsequent trials were correct based off of the error, they were only penalized for the first error.

Both N-Back tasks were 1-back tasks in which children needed to indicate whether a stimulus (auditory: robot band sound of 1,000Hz, 1,250Hz, 1500 Hz, 1,750 Hz, or 2,000 Hz or visual: domino-type pattern) matched the previous stimuli. After training demonstrating the procedure, children knew to press a button with a green sticker if the most recent stimulus was identical to the previous stimulus, or a red sticker if the stimulus was different. Each N-Back task had three blocks of 18 trials, which were evenly divided between stimuli that were identical or different.

Phonological Tasks—There were three phonological tasks: Digit Span, Digit Span Running, and Nonword Repetition. In Digit Span, a child repeated increasingly long series of single-syllable digits, in order. In Digit Span Running, a child repeated as many digits, in order, as s/he could remember from the end of an unpredictably long (7–10) string of numbers. In Nonword Repetition, a child repeated nonwords, increasing in length from 2 to 5 syllables.

Focus/Visuospatial Tasks—There were four Focus/Visuospatial tasks: Location Span, Location Span Running, Visual Span, Visual Span Running. In Location Span, children pointed to one of eight dots arranged in a circle to indicate where a series of arrows had pointed, with the series increasing in length across the task from 2 to 6 locations. Location Span Running was similar, except that children had to recall as many locations as possible, in order, from the end of a series of an unpredictably long (5–8) string of arrow locations. In Visual Span, children had to select, in order, the polygons they had seen presented from a field of polygons. The number of polygons increased as the task continued from 1 to 6 polygons. Visual Span Running was similar, except that children had to recall, in order, as many polygons from the end of a series of an unpredictably long (3–6) string of polygons.

Binding Tasks—There were three binding tasks: two within domain and one cross-domain. Phonological Binding (within domain) asked children to bind nonwords to non-speech sounds. This was presented in the context of robots ordering candy. After hearing the nonword and non-speech sounds presented together, children were prompted with the non-speech sound, and asked to produce the matching nonword. Visual-Spatial Binding (within domain) showed children polygons that were placed in a grid. Children were then asked to drag the polygons to the correct location in the grid. Finally, Cross-Modal binding (across domains) presented children with a polygon and a nonword. When children saw the polygon, they were asked to produce the corresponding nonword.

Statistical Analyses

In order to ensure that the cognitive structure of working memory did not differ for children with and without dyslexia, confirmatory factory analysis (CFA) was used to assess the fit of the combined model from Gray et al. (2017) for the group with dyslexia. Figure 2 shows the relationship between the hypothesized factors and the 13 task variables. All analyses were conducted with Mplus 8. 0 (Muthén & Muthén, 1998–2017) using maximum likelihood parameter estimation with standard errors and χ^2 test statistics that are relatively robust to non-normality (MLR). MLR allowed for missing data. Model fit was assessed globally using four statistics for the group with dyslexia: the $\chi 2$ test statistic, the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the N-adjusted Bayesian Information Criterion (N- adjusted BIC). Rejection of the null hypothesis based on the $\chi 2$ implies a lack of support for the hypothesized model. The CFI compares the fit of the hypothesized model to a null model and ranges in value from 0.00 to 1.00, with higher values indicating better fit. The traditional cutoff value for good fit with the CFI is .90 (Bentler, 1990). The RMSEA is an absolute index that yields the value of 0.00 if the model fits the data perfectly and increases in value with poorer fitting models. Browne and Cudeck (1992) suggested the following cutoffs for RMSEA: .10 or less for adequate fit, .08 or less for reasonable fit, and .05 or less for close fit. We also included the N-adjusted BIC. Based on this index, the model with the lower value demonstrates better fit. It penalizes for a lack of model parsimony taking into account sample sizes.

To be confident that it would be appropriate to compare the groups, we examined measurement invariance between the groups with dyslexia and TD (Thompson & Green, 2013) using the procedure outlined in Figure 1. We also examined reliability through the lens of the factor models. We (Gray et al., 2017, Table 3, p. 191) reported estimates and confidence intervals for the reliabilities for each of the working memory tasks for the sample of typically developing second graders, via split- half and spilt-third coefficients, both being special cases of the K-split coefficient (Green & Yang, 2015; Raju, 1977). Most reliabilities were moderate to high, with a range from .38 (Cross- Modal binding) to .95 (Number Updating). These are all within or above the range reported by other researchers evaluating individual reliabilities for working memory in children (e.g., Alloway et al., 2004; 2006; Bayliss et al., 2003; Hornung et al., 2011, Giofré et al., 2013; Michalczyk et al., 2013).

Evaluating the reliabilities for a working memory task separately is sensible if each such task would be used in isolation. In the present work, we model the tasks as indicators of constructs represented by latent factors. We therefore seek to understand the reliability for sets of tasks, grouped in terms of their associated factors. More specifically, we evaluate the maximal reliability for the set of tasks with respect to a factor (Bentler, 2007; Hancock & Mueller, 2001; Raykov, 2004) via Coefficient H, which is a function of the standardized loadings for the indicators of a factor (Hancock & Mueller, 2001). This is a measure of reliability of a factor, similar to the better-known coefficient alpha, but with each item contributing a different amount to the overall scale score based on its standardized factor loading (McNeish, 2018). In this work, we report Coefficient H for each factor based on its indicators, using the Excel spreadsheet provided by McNeish (2018).

Results

The means, standard deviations, and percent missingness for both groups for all tasks are reported in the supplemental materials (Table S1). Missing data resulted from factors unrelated to performance, such as non-systematic technological problems (e.g., a game froze) or timing issues (e.g., child had to leave data collection session early). Because missingness was limited and viewed as completely at random, it is unlikely to be a serious threat to the interpretation of the data.

We initially evaluated the combined working memory model from Gray et al. (2017) in the group with dyslexia, shown in Figure 2, because this was the best-fitting model for children with TD. Overall, global fit indices indicated that the model did not fit the data well, $\chi 2(59, N = 82) = 85.67$, $p = .013$; CFI = .75; RMSEA = .07, 90% CI [.04, .11], N-adjusted BIC = 3515.11. If the model did not fit, we could not claim invariance, and would not be justified in comparing the groups. So, we then evaluated modification indices to identify areas of the model that could be mis-specified. Modification indices suggested covariance between errors for the Visual Span Running and Digit Span Running tasks. After the addition of a correlated error term for these indicators, the revised model fit the data well, $\chi 2(58, N = 82) = 68.75$, $p = .15$; CFI = .90; RMSEA = .05, 90% CI [.00, .09], N-adjusted BIC = 3504.45. The parameter estimates for this model in the group with dyslexia are presented in Figure 3. Although the inclusion of the correlated error term was driven by the model fitting process rather than by theory, it improved model fit. The supplementary material includes a discussion of a potential theoretical consideration for this correlated error term. Keeping the basic structure of the best-fitting model for the children with TD and adding a correlated error term (a relatively minor adjustment) to achieve model fit and allow for between-group comparisons was preferable to abandoning the best-fitting model and doing exploratory analyses to look for other possibilities.

Assessment of Model Invariance

The first step in the assessment of measurement invariance across groups was establishing configural invariance: showing a model has the same factorial structure for both groups. If so, this implies the same patterns of correlations for both groups. The final model for the group with dyslexia was identical to the combined model proposed by Gray et al. (2017) with one revision: the addition of the covariance between the errors terms of the Digit Span Running and Visual Span Running tasks. Thus, to establish configural invariance, we assessed the fit of the revised combined model (i.e., the original model plus the correlated error term) in the group with TD. The model fit well for this group, establishing configural invariance, $\chi 2(58, N = 167) = 59.36$, $p = .43$; CFI = 1.00; RMSEA = .01, 90% CI [.00, .05], N-adjusted BIC = 7575.26. Figure 4 shows the parameter estimates of the revised combined model in the TD group.

We then conducted a multiple groups analysis. The fit for the models involved in assessing measurement invariance are presented in Table 2. The fit of these models is for the two groups combined. As shown, the fit for the revised combined model demonstrated good fit across groups when no between-group constraints were imposed. Thus, consistent with previous analyses of this model within each of the two groups, we concluded that the configuration for the two groups was sufficiently comparable (i.e., configural invariance) to proceed to the next step.

Next, to determine if the factor loadings were the same for each group, we assessed metric invariance in which between-group constraints were imposed on the factor loadings. If found invariant, this allows comparison of estimated factor variances and covariances across groups. Overall, as shown in Table 2, there was no decrement in fit when we imposed these constraints. The χ^2 difference test was nonsignificant, the RMSEAs, CFIs, and the N-adjusted BIC were slightly improved for the model with equality constraints between groups on the factor loadings. Accordingly, we concluded that the model was metrically invariant: factor loadings were not significantly different between the groups with and without dyslexia.

Next, to determine if the intercepts were the same between groups, we evaluated scalar invariance in which we imposed between-group constraints on both the factor loadings and the intercepts for the measurement tasks. If found to be invariant, this means that the latent variable means can be compared across groups. The model with scalar invariant constraints demonstrated slightly worse fit based on RMSEA and CFI, but somewhat better fit based on N-adjusted BIC; the χ^2 difference test evaluating invariance of intercepts was not significant. This provides evidence of scalar invariance: intercepts were not significantly different between the groups with and without dyslexia. As shown in Figure 1, we now had clearance to compare the groups on working memory performance.

Figure 5 presents select parameter estimates for the revised Gray et al. (2017) model for the groups with dyslexia and TD, under scalar invariance where the unstandardized loadings and intercepts are invariant. The supplemental materials include parameter estimates and standard errors for the unstandardized and standardized solutions for all parameters for this final measurement model of scalar invariance.

Reliability

Based on the scalar invariance model, the values of Coefficient H for the group with dyslexia were .58 for the Central Executive factor, .73 for the Focus/Visuospatial factor, and .50 for the Phonological factor. For the typically-developing group, the values were .70 for the Central Executive factor, .77 for the Focus/Visuospatial factor, and .53 for the Phonological factor. This suggests comparable reliability in the measurement of the factors across groups, save that the reliability for Central Executive is a bit higher in the typically-developing group than in the group with dyslexia. Though these values may be considered low in some circumstances, several points are noteworthy. First, there is no cutoff for what counts as adequate reliability that applies across all assessments and uses, as use depends on the targeted inference and stakes associated with that inference (AERA et al., 2014, Chapter 2). Second, these estimates are in line with those reported by other research on working memory in children which looked at the reliability of factors (e.g., Gathercole et al., 2004; Michalczyk et al., 2013). Finally, this notion of reliability is formulated at the level of inference about an individual. In our case, one of our foci is on the factor means for groups. As such, a more targeted notion of reliability would be at the level of the group mean, rather than the individual (AERA et al., 2014, Chapter 2). To our knowledge, procedures for evaluating reliability at the level of factor means and factor mean differences, rather than individuals, have not been developed.

Factor Mean Differences

To compare performance on working memory measures across groups, we examined factor mean differences (i.e., how children performed on each of the factors in the working memory model). As shown in Table 3, the means for the Central Executive, Focus/Visuospatial, and Phonological factors were significantly higher in the group with TD than in the group with dyslexia. Because of the way the factors are scaled, the difference in means and standard deviations on the factors are difficult to interpret. To obviate this difficulty, we computed an effect size statistic, d. The d statistic was computed by subtracting the mean for the group with dyslexia from the mean for the group with TD and dividing the result by the standard deviation of the group with TD. The pooled standard deviation was not used in the computation because of the heterogeneity of the variances between the two groups. On the Central Executive tasks and the Phonological factors, the d statistics of −1.00 and −.89, respectively, are relatively strong, whereas the d statistic of −.47 on the Focus /Visuospatial factor is moderate in value. Because the d statistic is computed for factors, we would expect the magnitudes to be greater than d statistics for observed measures.

Finally, we examined whether non-verbal intelligence and language could account for the differences in the factor means. To pursue this, we examined models which included nonverbal intelligence (K-ABC-2) and language (CELF-4) scores as predictors of the working memory factors in the multiple-group model, assuming scalar invariance. For the group with TD, the K-ABC-2 scores (mean of 117.60, standard deviation of 15.53) and CELF-4 scores (mean of 108.75, standard deviation of 9.59) had a correlation of .41. For the group with dyslexia, the K-ABC-2 scores (mean of 106.80, standard deviation of 13.40) and CELF-4 scores (mean of 100.30, standard deviation of 8.61) had a correlation of .25.

A series of models were fit. In each model, the K-ABC-2 and CELF-4 scores were permitted to have group-specific means, variances and covariances. (The models were initially fitted using the K-ABC-2 scores. However, the fitted models exhibited numerical instability, possibly due to the large variance associated with the K-ABC-2 scores. These scores were rescaled by dividing them by 10 before fitting the models; no evidence of numerical instability was reported.) The models differed regarding the assumptions of invariance of the regression structure across group. The most general model allowed for noninvariant intercepts, noninvariant slopes of the K-ABC-2 and CELF-4 scores, and a noninvariant error structure for the factors, including error covariances for the factors. The second model differed in that it constrained the error variances and error covariances to be invariant across groups. The third model additionally constrained the slopes of the K-ABC-2 and CELF-4 scores.

Table 4 presents the comparative fit of the models. The evidence from the χ^2 difference tests, RMSEA, CFI, and N-adjusted BIC supports the use of the last model, in that there is little loss of fit in assuming the slopes and error variances and covariances are invariant across groups.

In this model, for each factor the difference in group means is captured by the difference between intercepts. The intercepts for the group with TD was fixed to 0 to identify the model. The group difference is therefore captured by the intercept for the group with dyslexia. The estimated intercept for the Central Executive factor was −.04 (SE = .02, p = .01, 95% CI = (−.08, −.01)). The estimated intercept for the Focus/Visuospatial factor was −.04 (SE = .50, p = .94, 95% CI = (−1.01, .94)). The estimated intercept for the Phonological factor was −1.11 (SE = .48, p = .02, 95% CI = (−2.05, −.17)). These findings indicate that the difference between the group with TD and the group with dyslexia on the Focus/Visuospatial factor could be explained by differences on the K-ABC-2 and CELF-4. However, for the Central Executive factor and Phonological factor, group differences persist even after accounting for performance on the K-ABC-2 and CELF-4.

Discussion

Structure of Working Memory in Children with Dyslexia

The first purpose of this study was to determine whether a previously-tested working memory model that fit the data for second-grade children with TD (Gray et al. 2017) would be invariant when we tested the model with a group of second grade children with dyslexia. We established configural, metric, and scalar invariance. That is, we determined that the model was testing the same constructs for both groups, with similar factor loadings, and similar intercepts; using this model would allow us to compare apples to apples, as it were.

Once invariance was established, we compared factor means for the two groups. To our knowledge, this is the first study to establish working memory model invariance across a group of children with TD and dyslexia. Often, researchers compare group means without knowing whether the structure of the constructs they are testing is invariant across groups. The fact that we have established invariance increases the validity of our between-group findings.

Finding invariance is not surprising, given that there are no theories suggesting that working memory factors should differ in children with or without dyslexia. While work from D'Mello and Gabrieli (2018) offered the potential for brain differences leading to functional invariance, overall our finding is consistent with the nature of the neurodiversity they outlined between children with dyslexia and peers with TD. Specifically, they found differences in quantity of activation or density of structures rather than differences in brain regions utilized during tasks.

Because we established invariance, it was acceptable to compare group means. In our sample, we found lower factor means for children with dyslexia compared to their peers with TD. The idea that children with dyslexia are working within the same constructs of working memory as their unimpaired peers, but are more challenged to score well on tasks in different domains of working memory, is mirrored in other findings about how children with dyslexia learn. For example, in the area of phonology, Baron et al. (2018) found that children with dyslexia benefitted from orthographic facilitation; they improved their learning when a word was presented with orthography compared to an auditory-only presentation, like their typically-developing peers. However, the children with dyslexia did not gain as much benefit as their peers. They appeared to be tapping into the same system for learning but were not as effective.

Working Memory Performance in Children with Dyslexia

The significant differences in factor group means that remained after including language and nonverbal intelligence in the model were in two factors: Phonology and the Central Executive. The between-group difference in the Focus/Visuospatial factor was no longer significant. This finding was not unexpected given the strong relation between nonverbal intelligence and the Focus/Visuospatial factor found in Gray et al., (2017). Gray et al. (2017) examined the relationship between working memory factors and nonverbal intelligence factors in both the Cowan and Baddeley and Hitch models. Both nonverbal intelligence factors (fluid intelligence and visual processing intelligence) were most strongly related to the Focus of Attention (Cowan) and Visuospatial factors (Baddeley & Hitch) in each of these models, with correlations above .6 for each model.

Lower scores on the Phonological factor for the dyslexic group relative to the TD group were expected. Phonological deficits are a core feature of dyslexia at the group level (e.g., Vellutino et al., 2004) and, specifically, verbal or phonological working memory deficits have been reported by many teams (e.g., Alloway et al., 2017; Berninger et al., 2006; 2008) using a wide variety of tasks. Our study provided a robust test of group differences in phonological working memory because we tested these differences using factor scores, which pool shared variance on multiple tests of phonology.

The difference between groups for the Central Executive factor, even after controlling for language and nonverbal intelligence, was less expected, although not unsupported by the literature. For example, Berninger et al. (2008) identified Central Executive deficits in some participants in their sample, but reported these more in the context of individual differences (via a mixed modelling approach), rather than a group effect. Recall that Gray et al. (2019) also found individual differences in the working memory profiles of children with dyslexia, with some children with dyslexia demonstrating Central Executive deficits while other did not. Schuchardt et al. (2008) also identified Central Executive deficits in children with dyslexia, but in contrast to our findings, their effect disappeared when they controlled for phonological working memory skills. On the other hand, Reiter et al. (2005) found evidence for Central Executive deficits, including those central to working memory, in children with dyslexia as did Varvara et al. (2014). Swanson (1993) has suggested that ‘executive processing’ is related to reading comprehension deficits in children with reading disabilities and Varvara et al. (2014) also point to auditory and visual-spatial attention as explaining some of the variance in reading performance in children. Brandenburg and colleagues (2015) also found that Central Executive deficits were related to reading disabilities in 3rd graders

Potential difference between our findings (i.e., no group differences in the Focus of Attention/Visuospatial factor between groups) and some past literature are likely due to the tasks we used and the fact that we accounted for nonverbal intelligence and language scores. We used the CABC-WM because it intentionally limited the verbal contributions of the visual tasks (e.g., difficult-to-name polygons, dots placed at angles that are not easily tied to clock points). Alternatively, many other tasks that tap into Focus/Visuospatial skills have strong linguistic and phonological components. For example, any task that includes real orthography will tap into phonology for a child with reading skills (e.g., Ricketts, Bishop, & Nation, 2009). Thus, studies that find differences in visuospatial skills may actually be detecting well-established deficits in phonology that are inherent in their tasks. A second explanation is that not all studies that examine Focus/Visuospatial working memory examine the role of nonverbal intelligence or language. Our initial group difference was rendered nonsignificant when considering oral language and nonverbal intelligence. Had we not explored the role of nonverbal intelligence or language, we would be making different statements about Focus/Visuospatial skills in children with dyslexia.

Importantly, we note that the model presented above represents one possible structure of the relationships among working memory factors, non-verbal intelligence, and language. We use these models to pursue our questions regarding whether group differences in working memory may be related to group differences in the non-verbal intelligence and language, not as claim on the causal structure among these aspects of cognition. This and other possible structures are deserving of attention and comparison in future theoretical and empirical research.

Another potential group difference worth commenting on is the strength of the cross-task correlation between Visual and Digit Running spans. This resembles the group difference in the strength of the correlation between the Central Executive and Phonological factors (.63 for the group with dyslexia; .17 for the typically-developing group). For these two factors, recall that a working memory difference between groups persisted after variance from language and intelligence tests was accounted for. Similarly, Berninger and colleagues (2006) found an unusually high correlation between executive processes and phonological processes in their adults with severe dyslexia. Perhaps impairment in the Central Executive factor drives deficiency in Phonological working memory skills and also plays a formally unacknowledged role in our cross-correlated running span tasks.

Practical Implications

Prior to drawing conclusions about working memory deficits in groups of children with dyslexia relative to their peers with TD, it is important to have evidence of configural, metric, and scalar invariance across those groups. Without this evidence, one should not compare mean differences between groups because one cannot interpret the validity of the results. This study is the first to meet these standards to examine group differences in working memory in children with and without dyslexia.

Our finding of Phonological working memory deficits was expected, but the finding of Central Executive deficits in children with dyslexia means that educators need to pay attention to the Central Executive demands of educational tasks. Working memory is not amenable to training (e.g., Melby-Lervåg & Hulme, 2013). However, if educators are aware of the Central Executive demands of a task (e.g., the need to update information about a pronoun’s referent while reading a sentence for comprehension), they can provide learners with compensatory supports to help alleviate Central Executive deficits. Although we did not find evidence for deficits for children with dyslexia in the Focus/Visuospatial factor in terms of working memory after controlling for nonverbal intelligence and language, children with dyslexia did perform more poorly on these tasks. Thus, children with dyslexia with lower nonverbal IQs and lower language skills (note that all the children in this study had normal intelligence and language skills) might be more likely to struggle with learning that relies on visual and spatial information. For example, including pictures as memory cues might not work equally well for all children with dyslexia.

Limitations

Although our participant group was free from the interpretive challenges posed by studies that have children with dyslexia and other common comorbid conditions (i.e., we ruled out those with developmental language disorder), we did not rule out dyscalculia, which can make unique contributions to working memory. However, Schuchardt et al. (2008) found that children with dyscalculia have issues with visuospatial working memory, which was a relative strength of our cohort.

A broader point of caution is that our findings of invariance may be, in part, due to the need for an even larger group of children with dyslexia. To our knowledge, this is the first study to examine the possible invariance in working memory between children with typical development and those with dyslexia. Future studies would benefit from pursuing questions of invariance in larger samples.

Future Directions

The results of tests for configural, metric, and scalar invariance provide support for the notion of measurement invariance sufficient to pursue the group comparisons on the latent factors. However, to achieve model fit, we added a correlated error term between the Digit Span Running and Visual Span Running tasks (in the models for separate groups, this was estimated to be .43 for the group with dyslexia, .16 for the typically-developing group). At this point it is not yet clear why this was needed. One possibility is it is a result of shared variance due to the method of measurement. This possibility was strengthened by the fact that the two measures that had correlated error were running tasks. Cowan et al. (2017) found deficits for children with dyslexia relative to running tasks, but the only deficit that was specific to running tasks (and not regular span tasks as well), was for Location Span Running - the one running task that was not indicated as part of the correlated error term. Thus, an explanation based on the 'running' nature of the tasks is not, at present, a satisfactory explanation. Future research validating the structure of working memory in children with dyslexia is needed to ensure that our structure is again obtained, including the details that go beyond what Gray et al. (2017) obtained with typically-developing children.

Though the test of metric invariance (i.e., factor loadings) supported the hypothesis of invariance, in the configural invariance model (i.e., constructs) there were differences between the groups on some the loadings. One of these was the loading for Digit Span Running along the Focus of Attention/Visuospatial Sketchpad. This standardized loading for the group with dyslexia (−.06) was weaker than for the typically developing group (.28).

Further high-powered investigation may show that, in subtle ways, scalar invariance does not completely hold after all because of slightly different ways in which the tasks are approached by children with typical development versus those with dyslexia. In our combined model with scalar invariance (Figure 5), the standardized correlated error term between visual and digit running span was significant for children with dyslexia (.37) but not typically-developing children (.18). The only other notable difference was in the standardized correlation between the central executive and phonological factors, significant for children with dyslexia (.47) but not for typically-developing children (.17). These correlations may indicate that the phonological processes that are involved in running digit span and other phonological tasks require more executive function for children with dyslexia, a hypothesis that could lead to research on the role of effort directed at the phonological storage aspects of reading in the two groups.

Summary

Establishing invariance for this model of working memory for these two groups of second-grade children provides researchers with evidence that the structure, factor loadings, and intercepts of working memory do not vary for children with TD and dyslexia. Knowing this, we can be more confident that our tests of between-group findings are valid. We have found, after accounting for language and non-verbal intelligence skills, that children with dyslexia have lower performance on Phonological and Central Executive skills. Differences in nonverbal intelligence and language accounted for differences in Focus/Visuospatial skills. While this information helps us understand more about dyslexia at the group level, future work will need to identify the role of individual differences so that individualized intervention strategies can be developed. Additional research is necessary to directly examine the relation between working memory, learning, and academic achievement. However, establishing invariance in a model of working memory between children with and without dyslexia is a major step toward understanding working memory differences between children with dyslexia and TD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We are deeply grateful to the staff, research associates, school administrators, teachers, children, and families who participated. Key personnel included (in alphabetical order) Shara Brinkley, Katy Cabbage, Margeaux Ciraolo, Cecilia Figueroa, Karen Guilmette, Trudy Kuo, Bjorg LeSueur, Annelise Pesch, Anthony Resnick, and Jean Zimmer. Many students also contributed to this work including (in alphabetical order) Genesis Arizmendi, Lauren Baron, Alexander Brown, Gary Carstensen, Nora Schlesinger, Nisha Talanki, and Hui-Chun Yang. We gratefully acknowledge the important contributions of our late colleague and co-investigator Sam Green in the conceptualization and design of the research leading to this article.

Sponsors: This work was supported by funding from the National Institutes of Health NIDCD Grant #R01 DC010784.

References

AERA, APA, & NCME. (2014). Standards for educational and psychological testing. American Educational Research Association.

Alloway TP, Gathercole SE, & Pickering SJ (2006). Verbal and visuo-spatial short- term memory and working memory in children: Are they separable? Child Development, 77, 1698–1716. 10.1111/j.1467-8624.2006.00968.x [PubMed: 17107455]

Alloway TP, Gathercole SE, Willis C, & Adams AM (2004). A structural analysis of working memory and related cognitive skills in young children. Journal of Experimental Child Psychology, 87, 85–106. 10.1016/j.jecp.2003.10.002 [PubMed: 14757066]

Alloway TP, Tewolde F, Skipper D, & Hijar D (2017). Can you spell dyslexia without SLI? Comparing the cognitive profiles of dyslexia and specific language impairment and their roles in learning. Research in Developmental Disabilities, 65, 97–102. 10.1016/j.ridd.2017.04.013 [PubMed: 28486125]

Alt M, Hogan T, Green S, Gray S, Cabbage K, & Cowan N (2017). Word learning deficits in children with dyslexia. Journal of Speech, Language, and Hearing Research, 60, 1012–1028. 10.1044/JSLHR-L-16-0036

Baddeley A (2000). The episodic buffer: A new component of working memory? Trends in Cognitive Sciences, 4, 417–423. 10.1016/S1364-6613(00)01538-2 [PubMed: 11058819]

Baddeley AD, & Hitch G (1974). Working memory. In Bower GA (Ed.) Recent advances in learning and motivation (Vol. 8, pp. 47–90). Academic Press.

Baddeley A, Hitch G, & Allen R (2021). A multicomponent model of working memory. In Logie RH, Camos V, and Cowan N (eds.), Working memory: State of the science. Oxford University Press. (pp. 10–43)

Baron LS, Hogan TP, Alt M, Gray S, Cabbage KL, Green S, & Cowan N (2018). Children with dyslexia benefit from orthographic facilitation during spoken word learning. Journal of Speech, Language, and Hearing Research, 61, 2002–2014. 10.1044/2018_JSLHR-L-17-0336

Bayliss DM, Jarrold C, Gunn DM, & Baddeley AD (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. Journal of Experimental Psychology: General, 132(1), 71–92. 10.1037/0096-3445.132.1.71 [PubMed: 12656298]

Bentler PM (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107, 238–246. [PubMed: 2320703]

Bentler PM (2007). Covariance structure models for maximal reliability of unit-weighted composites. In Lee S (Ed.), Handbook of computing and statistics with applications: Vol. 1. (pp. 1–19). New York, NY: Elsevier.

Berninger VW, Abbott RD, Thomson J, Wagner R, Swanson HL, Wijsman EM, & Raskind W (2006). Modeling phonological core deficits within a working memory architecture in children and adults with developmental dyslexia. Scientific Studies of Reading, 10, 165–198. 10.1207/s1532799xssr1002_3

Berninger VW, Raskind W, Richards T, Abbott R, & Stock P (2008). A multidisciplinary approach to understanding developmental dyslexia within working- memory architecture: Genotypes, phenotypes, brain, and instruction. Developmental Neuropsychology, 33, 707–744. 10.1080/87565640802418662 [PubMed: 19005912]

Brandenburg J, Klesczewski J, Fischbach A, Schuchardt K, Büttner G, & Hasselhorn M (2015). Working memory in children with learning disabilities in reading versus spelling: Searching for overlapping and specific cognitive factors. Journal of Learning Disabilities, 48(6), 622–634. 10.1177/0022219414521665 [PubMed: 24532828]

Browne MW, & Cudeck R (1992). Alternative ways of assessing model fit. Sociological Methods & Research, 21, 230–258. 10.1177/0049124192021002005

Cabbage KL, Brinkley S, Gray S, Alt. M, Cowan N, Green S, Kuo T, & Hogan T (2017). Assessing working memory in children: The Comprehensive Assessment Battery for Children – Working Memory (CABC-WM). Journal of Visualized Experiments (JoVE), 124, e55121.

Cain K, Oakhill J, & Bryant P (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. Journal of Educational Psychology, 96, 31–42. 10.1037/0022-0663.96.1.31

Campos IS, Almeida LS, Ferreira AI, & Martinez LF (2013). Working memory as separable subsystems: A study with Portuguese primary school children. Spanish Journal of Psychology, 16, e14, 1–10. 10.1017/sjp.2013.6

Centanni TM (2020). Neural and genetic mechanisms of dyslexia. In: Argyeopoulous GPD (Ed.) Translational Neuroscience of Speech and Language Disorders. (pp. 47–68). Springer. 10.1007/978-3-030-35687-3_4

Cowan N (1995a). Attention and memory: An integrated framework. Oxford University Press.

Cowan N (1999). An embedded-processes model of working memory. In Miyake A & Shah P (Eds.), Model of working memory: Mechanism of active maintenance and executive control (pp. 62–101). Cambridge University Press

Cowan N (2001). The magical number four in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences, 24, 87–114. 10.1017/S0140525X01003922

Cowan N (2014). Working memory underpins cognitive development, learning, and education. Educational Psychology Review, 26, 197–223. DOI: 10.1007/s10648-013-9246-y. [PubMed: 25346585]

Cowan N (2017). The many faces of working memory and short-term storage. Psychonomic Bulletin & Review, 24(4), 1158–1170. 10.3758/s13423-016-1191-6 [PubMed: 27896630]

Cowan N, Hogan TP, Alt M, Green S, Cabbage KL, Brinkley S, & Gray S (2017). Short-term memory in childhood dyslexia: Deficient serial order in multiple modalities. Dyslexia, 23, 209–233. 10.1002/dys.1557 [PubMed: 28497530]

D'Mello AM, & Gabrieli JD (2018). Cognitive neuroscience of dyslexia. Language, Speech, and Hearing Services in Schools, 49, 798–809. 10.1044/2018_LSHSS-DYSLC-18-0020

Daneman M, & Carpenter PA (1980). Individual differences in working memory and reading. Journal of Memory and Language, 19, 450–466.

Dollaghan C, & Campbell TF (1998). Nonword repetition and child language impairment. Journal of Speech, Language, and Hearing Research, 41, 1136–1146. 10.1044/jslhr.4105.1136

DuPaul GJ, Anastopoulos AD, Power TJ, Reid R, Ikeda MJ, & McGoey KE (1998). Parent ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. Journal of Psychopathology and Behavioral Assessment, 20, 83–102.

Fisher R, & Karl JA (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. Frontiers in Psychology, 10:1057. 10.3389/fpsyg.2019.01507 [PubMed: 31143151]

Gathercole SE, Pickering SJ, Ambridge B, & Wearing H (2004). The structure of working memory from 4 to 15 years of age. Developmental Psychology, 40(2), 177. 10.1037/0012-1649.40.2.177 [PubMed: 14979759]

Giofrè D, Mammarella IC, & Cornoldi C (2013). The structure of working memory and how it relates to intelligence in children. Intelligence, 41(5), 396–406. 10.1016/j.intell.2013.06.006

Goldman R & Fristoe M (2000). Goldman-Fristoe Test of Articulation – 2. Circle Pines, MN: American Guidance Service, Inc.

Gray S, Fox AB, Green S, Alt M, Hogan TP, Petscher Y, & Cowan N (2019). Working memory profiles of children with dyslexia, developmental language disorder, or both. Journal of Speech, Language, and Hearing Research, 62, 1839–1857. 10.1044/2019_JSLHR-L-18-0148

Gray S, Green S, Alt M, Hogan T, Kuo T, Brinkley S, & Cowan N (2017). The structure of working memory in school-age children and its relationship to intelligence. Journal of Memory and Language, 92, 183–201. 10.1016/j.jml.2016.06.004 [PubMed: 27990060]

Green SB, & Yang Y (2015). Evaluation of dimensionality in the assessment of internal consistency reliability: Coefficient alpha and omega coefficients. Educational Measurement: Issues and Practice, 34, 14–20. 10.1111/emip.12100

Hancock GR, & Mueller RO (2001). Rethinking construct reliability within latent variable systems. In Cudeck R, du Toit S, & Sörbom D (Eds.), Structural equation modeling: Present and future —A festschrift in honor of Karl Jöreskog (pp. 195–216). Lincolnwood, IL: Scientific Software International.

Hornung C, Brunner M, Reuter RA, & Martin R (2011). Children’s working memory: Its structure and relationship to fluid intelligence. Intelligence, 39(4), 210–221. 10.1016/j.intell.2011.03.002

Jeffries S, & Everatt J (2004). Working memory: Its role in dyslexia and other specific learning difficulties. Dyslexia, 10(3), 196–214. 10.1002/dys.278 [PubMed: 15341198]

Kaufman AS & Kaufman NL (2004). Kaufman Assessment Battery for Children, Second Edition. Circle Pines, MN: American Guidance Service, Inc.

Knoop-van Campen CA, Segers E, & Verhoeven L (2018). How phonological awareness mediates the relation between working memory and word reading efficiency in children with dyslexia. Dyslexia, 24, 156–169. 10.1002/dys.1583 [PubMed: 29577521]

Litt RA, Wang H-C, Sailah J, Badcock NA, & Castles A (2019). Paired associate learning deficits in poor readers: The contribution of phonological input and output processes. Quarterly Journal of Experimental Psychology, 72, 616–633. 10.1177/1747021818762669

Maehler C, & Schuchardt K (2016a). Working memory in children with specific learning disorders and/or attention deficits. Learning and Individual Differences, 49, 341–347. 10.1016/j.lindif.2016.05.007

Maehler C, & Schuchardt K (2016b). The importance of working memory for school achievement in primary school children with intellectual or learning disabilities. Research in Developmental Disabilities, 58, 1–8. 10.1016/j.ridd.2016.08.007 [PubMed: 27567244]

McNeish D (2018). Thanks coefficient alpha, we’ll take it from here. Psychological Methods, 23(3), 412–433. 10.1037/met0000144 [PubMed: 28557467]

Melby-Lervåg M, & Hulme C (2013). Is working memory training effective? A meta-analytic review. Developmental Psychology, 49, 270–291. 10.1037/a0028228 [PubMed: 22612437]

Menghini D, Finzi A, Carlesimo GA, & Vicari S (2011). Working memory impairment in children with developmental dyslexia: Is it just a phonological deficit? Developmental Neuropsychology, 36, 199–213. 10.1080/87565641.2010.549868 [PubMed: 21347921]

Michalczyk K, Malstädt N, Worgt M, Könen T, & Hasselhorn M (2013). Age differences and measurement invariance of working memory in 5-to 12-year-old children. European Journal of Psychological Assessment, 29, 220–229, 10.1027/1015-5759/a000149

Muthén LK & Muthén BO (1998–2017). Mplus user's guide (8th ed.). Los Angeles, CA: Muthén & Muthén.$^{\text{th}}$

Nadler RT, & Archibald LMD (2014). The assessment of verbal and visuospatial working memory with school age Canadian children. Canadian Journal of Speech-Language Pathology and Audiology, 38(3), 262–279.

Peng P, Barnes M, Wang C, Wang W, Li S, Swanson HL, Dardick W, & Tao S (2018). A meta-analysis on the relation between reading and working memory. Psychological Bulletin, 144, 48–76. 10.1037/bu10000124 [PubMed: 29083201]

Peng P, Namkung J, Barnes M, & Sun C (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. Journal of Educational Psychology, 108, 455. 10.1037/edu0000079

Putnick DL, & Bornstein MH (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review, 41, 71–90. Doi: 10.1016/j.dr.2016.06.004 [PubMed: 27942093]

Raju NS (1977). A generalization of coefficient alpha. Psychometrika, 42, 549–565.

Raykov T (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. Behavior Therapy, 35, 299–331.

Reiter A, Tucha O, & Lange KW (2005). Executive functions in children with dyslexia. Dyslexia, 11, 116–131. 10.1002/dys.289 [PubMed: 15918370]

Rhodes SM, Booth JN, Palmer LE, Blythe RA, Delibegovic M, & Wheate NJ (2016). Executive functions predict conceptual learning of science. British Journal of Developmental Psychology, 34, 261–275. 10.1111/bjdp.12129

Ricketts J, Bishop DVM and Nation K (2009). Orthographic facilitation in oral vocabulary acquisition. Quarterly Journal of Experimental Psychology, 62, 1948–1966. 10.1080/17470210802696104

Satorra A, & Bentler PM (2010). Ensuring positiveness of the scaled difference Chi-square test statistic. Psychometrika, 75, 243–248. [PubMed: 20640194]

Schuchardt K, Maehler C, & Hasselhorn M (2008). Working memory deficits in children with specific learning disorders. Journal of Learning Disabilities, 41, 514–523. 10.1177/0022219408317856 [PubMed: 18625783]

Semel E, Wiig EH, & Secord WA (2003). Clinical Evaluation of Language Fundamentals – Fourth Edition. Circle Pines, MN: American Guidance Service, Inc.

Swanson HL (1993). Executive processing in learning-disabled readers. Intelligence, 17(2), 117–149. 10.1016/0160-2896(93)90024-Y

Swanson HL, & Sachse-Lee C (2001). A subgroup analysis of working memory in children with reading disabilities: Domain-general or domain-specific deficiency?. Journal of Learning Disabilities, 34, 249–263. 10.1177/002221940103400305 [PubMed: 15499879]

Thompson MS, & Green SB (2013). Evaluating between-group differences in latent variable means. In Hancock GR & Mueller RO (Eds.), A Second Course in Structural Equation Modeling. 2nd Ed. (pp. 163–218). Greenwich, CT: Information Age Publishing.$^{\text{nd}}$

Toffalini E, Tomasi E, Albano D, & Cornoldi C (2018). The effects of the constancy of location and order in working memory visual–phonological binding of children with dyslexia. Child Neuropsychology, 24, 671–685. 10.1080/09297049.2017.1329411 [PubMed: 28532302]

Torgesen J, Wagner R, & Rashotte C (2012). Test of Word Reading Efficiency, Second Edition. Circle Pines, MN: American Guidance Service, Inc.

Varvara P, Varuzza C, Padovano Sorrentino AC, Vicari S, & Menghini D (2014). Executive functions in developmental dyslexia. Frontiers in Human Neuroscience, 8, 120. 10.3389/fnhum.2014.00120 [PubMed: 24639640]

Vellutino FR, Fletcher JM, Snowling MJ, & Scanlon DM (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? Journal of Child Psychology and Psychiatry, 45, 2–40. 10.1046/j.0021-9630.2003.00305.X [PubMed: 14959801]

Williams KT (2007). Expressive Vocabulary Test (2nd ed.). San Antonio, TX.: Pearson.$^{\text{nd}}$

Wong AMY, Ho CSH, Au TKF, McBride C, Ng AKH, Yip LPW, & Lam CCC (2017). Reading comprehension, working memory and higher-level language skills in children with SLI and/or dyslexia. Reading and Writing, 30, 337–361.

Woodcock R (2011). Woodcock Reading Mastery Test–Third Edition. Bloomington, MN: Pearson

Yoon M, & Lai MHC (2018). Testing factorial invariance with unbalanced samples. Structural Equation Modeling: A Multidisciplinary Journal, 25, 201–203. 10.1080/10705511.2017.1387859

Practitioner Points:

- We should not assume models describing how cognitive processes work for typically-developing children are the same for children with dyslexia.
- Our testing shows that one statistical model of working memory does fit both children with dyslexia and those with typical development.
- The finding suggests that the two groups carry out the working memory tasks in similar ways.
- After accounting for language and non-verbal intelligence skills, children with dyslexia had lower performance on Phonological and Central Executive skills.
- Additional research is necessary to directly examine the relation between working memory, learning, and academic achievement.

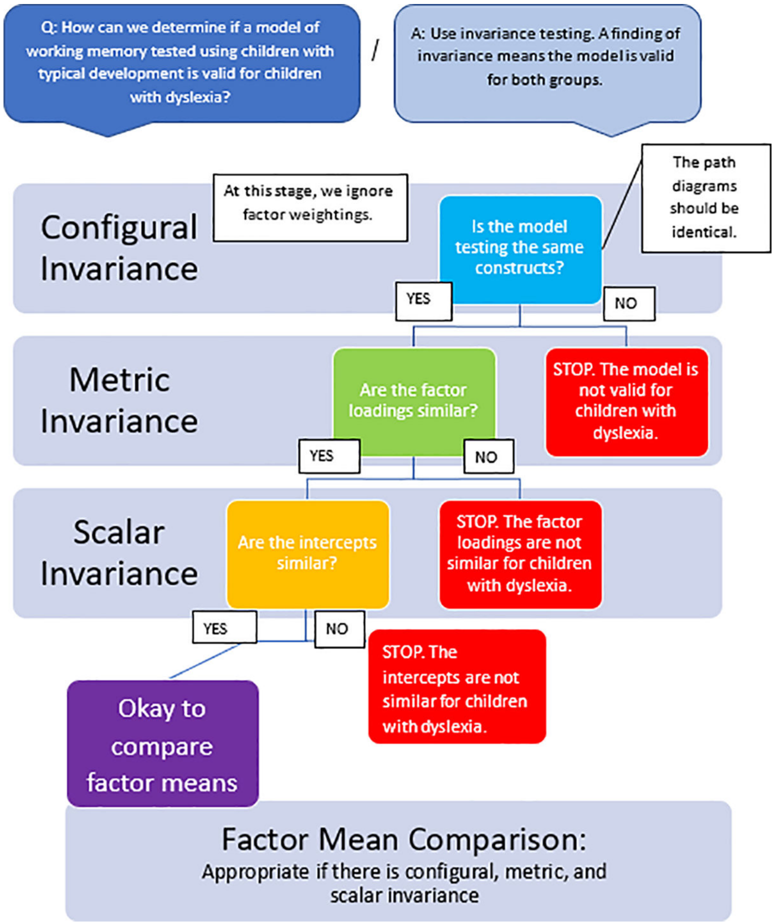

Figure 1.

Process of establishing invariance and justification for testing between-group means.

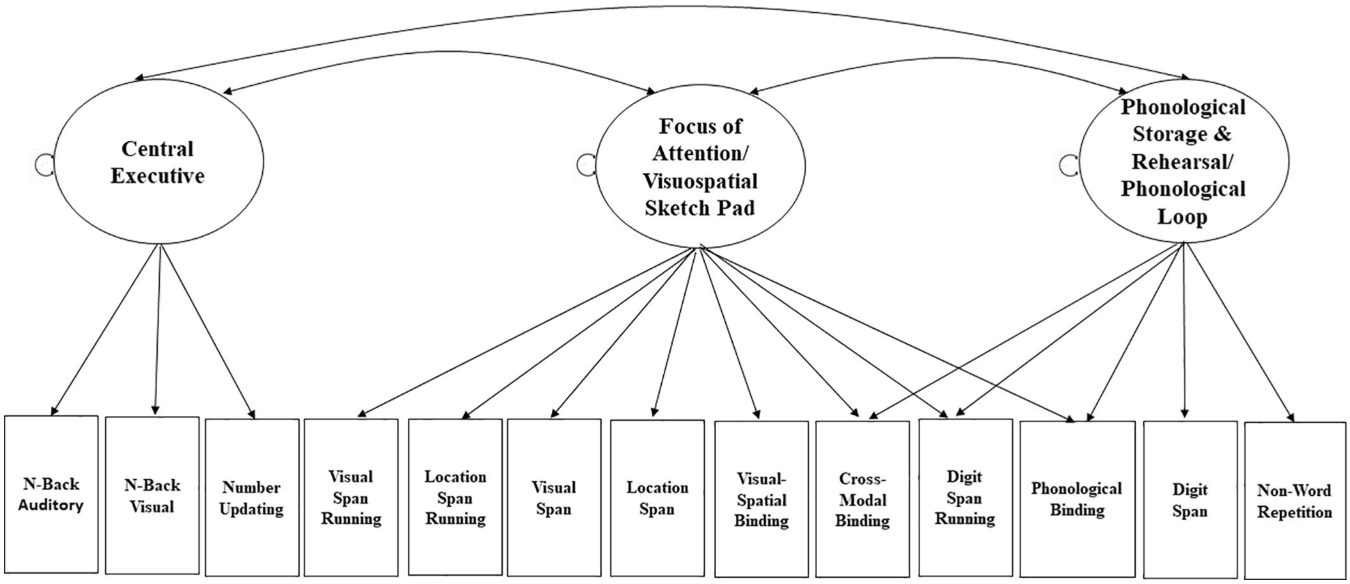

Figure 2.

Gray et al. (2017) Combined Model of Working Memory

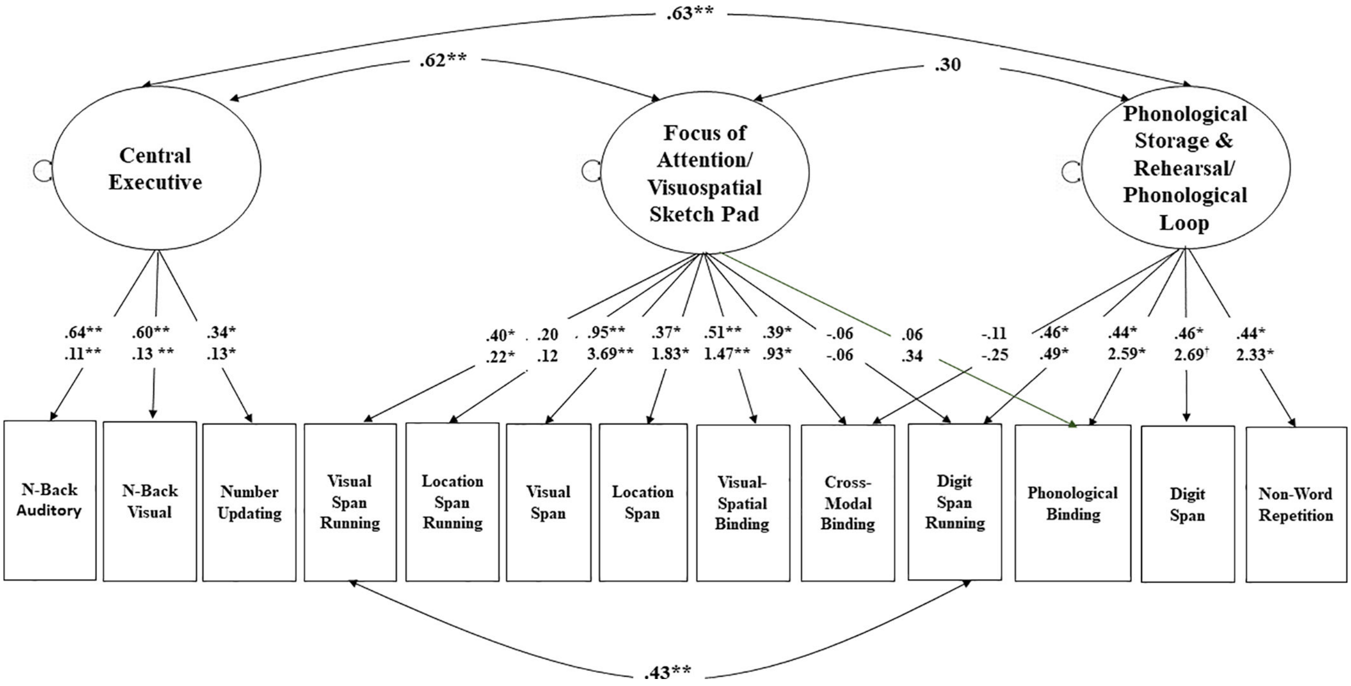

FIGURE 3:

CONFIGURAL INVARIANCE STAGE: Revised combined working memory model for children with dyslexia with standardized and unstandardized estimates presented first and second, respectively.

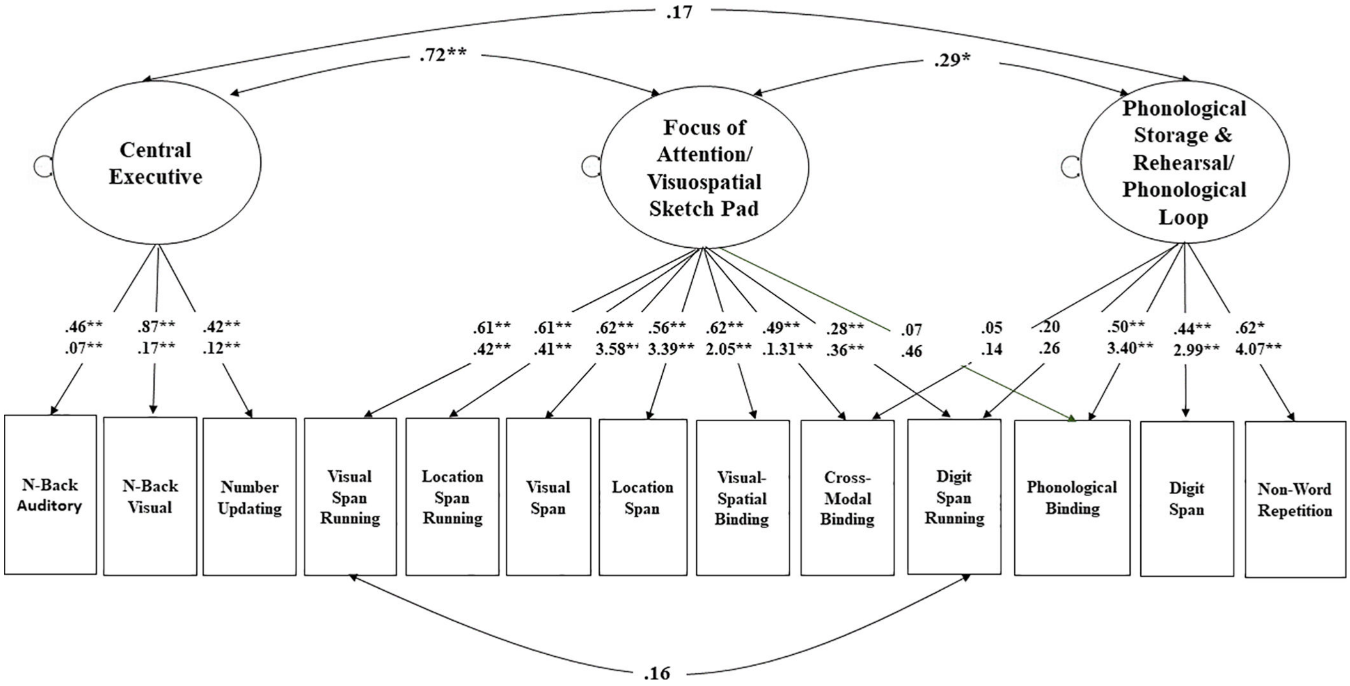

FIGURE 4:

CONFIGURAL INVARIANCE STAGE: Revised combined working memory model for typically-developing children with standardized and unstandardized estimated presented first and second, respectively.

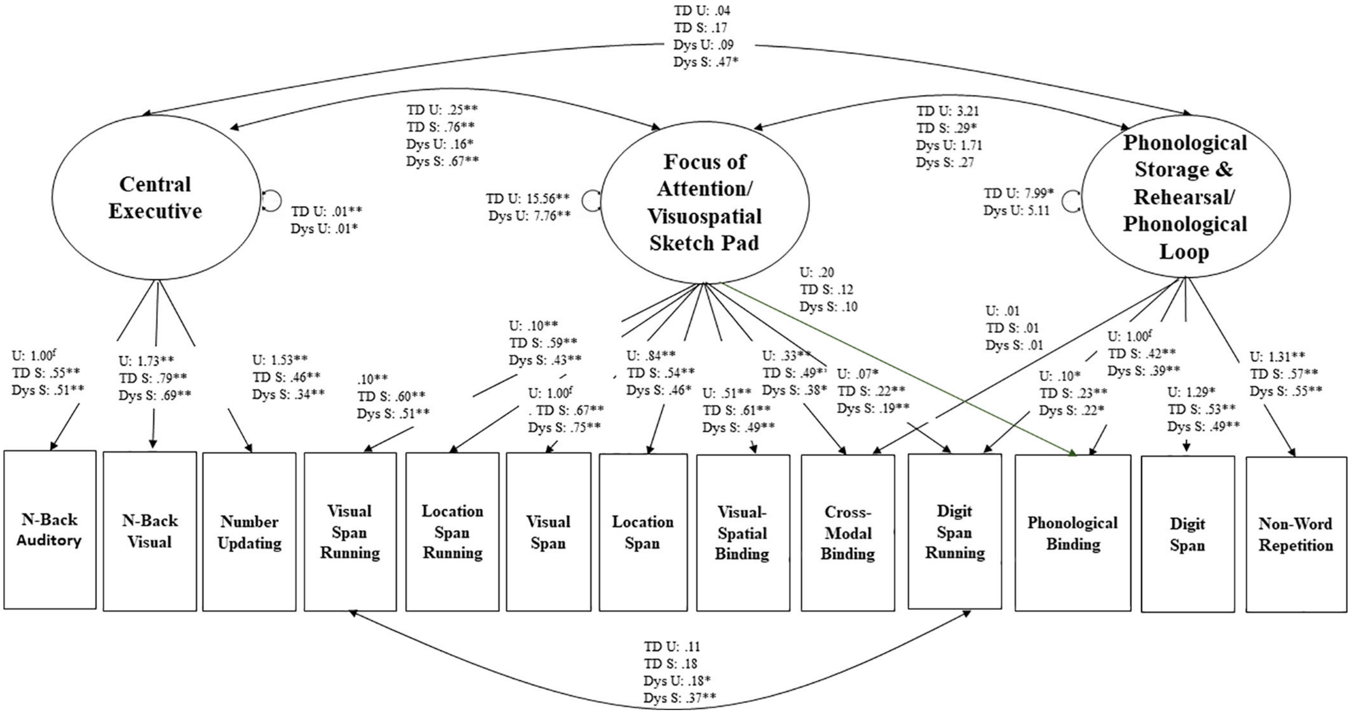

FIGURE 5:

Parameter estimated for the revised combined model for typically-developing children and children with dyslexia from the scalar invariance model. U = Unstandardized estimate. S = Standardized estimate. TD = Typically-developing. Dys = Dyslexic. * p < .05. ** p < .001. f = Fixed.

Table 1.

Participant information for descriptive and inclusionary measures

	TD (N = 167)			Dyslexia (N = 82)		
Measure	***M***	***SD***	**Range**	***M***	***SD***	**Range**
Age (years; months)	7;8	0;5	6;8–9;2	7;8	0;5	7;1–8;11
Mother's education in years *	15.39	1.66	12–17	14.84	1.91	9–17
GFTA-2 articulation accuracy percentile *	50.89	8.54	7–62	41.37	15.97	3–62
CELF-4 core language standard score *	108.75	9.59	88–130	100.30	8.61	88–126
TOWRE-2 composite standard score *	109.45	8.40	96–145	80.61	6.23	55–88
K-ABC 2 nonverbal index standard score *	117.60	15.53	78–160	106.80	13.40	82–141
EVT-2 standard score *	112.39	10.95	90–137	103.34	10.85	79–137
WRMT-III standard score *	108.23	9.85	82–144	93.89	10.66	55–115
NWR % percent consonants correct *	85.78	9.39	37.5–100	81.72	10.74	35.71–100
ADHD Rating Scale-IV Home Version *	10.19	8.77	0–41	13.02	9.02	0–40

* Indicates a group difference at $p < .05$

TD= typical development; GFTA-2 = Goldman-Fristoe Test of Articulation-2 (Goldman & Fristoe, 2000); CELF-4 = Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel, Wiig, & Secord, 2003); TOWRE-2 = Test of Word Reading Efficiency–Second Edition (Torgesen, Wagner, & Rashotte, 2012); K-ABC-2 = Kaufman Assessment Battery for Children–Second Edition (Kaufman &Kaufman, 2004); EVT-2 = Expressive Vocabulary Test–Second Edition (Williams, 2007); WRMT = Woodcock Reading Mastery Test- Third Edition, Paragraph Comprehension Subtest (Woodcock, 2011); NWR % CC = Nonword Repetition task, percent consonants correct (Dollaghan & Campbell, 1998). ADHD = parental rating of attention-defecit/hyperactivity disorder behaviors using the ADHD Rating Scale-IV Home Version (DuPaul et al., 1998).

Table 2.

Fit of Configural, Metric, and Scalar Invariance Models

MI Models	Model χ^2	χ^2 Comparison of Previous and Current Models[1]	RMSEA and 90% CI	CFI	*N*-adjusted BIC
Configural Invariance	χ^2 (116) = 128.07, p = .021	--	.029 [.00, .60]	.97	11148.25
Metric Invariance	χ^2 (129) = 140.50, p = .23	χ^2 (13) = 12.80 p = .46	.027 [.00, .05]	.97	11133.42
Scalar Invariance	χ^2(139) = 155.07, p = .17	χ^2 (10) = 14.44, p = .15	.030 [.00, .05]	.96	11124.73

[1] This chi square difference test was conducted taking into account that the model χ^2 were based on maximum likelihood robust estimation (Satorra & Bentler, 2010).

Table 3.

Difference in Means between Groups with Typical Development (TD) and Dyslexia

Latent Factor	TD Mean[1] (SD)	Dyslexic Mean (SD)	*d* Statistic[2]	*p* for Difference in Means
Central Executive	0 (.08)	−.08 (.08)	1.00	< .001
Focus of Attention/Visuospatial	0 (3.94)	−1.80 (2.78)	0.47	< .001
Pkonological Awareness/Phonological	0 (2.83)	−2.52 (2.26)	0.89	< .001

[1] The factor means were fixed to zeros for model identification.

[2] The *d* statistic was computed by subtracting the two factor means and dividing by the factor standard deviation of the TD group. The pooled standard deviation was not used in the computation because of the heterogeneity of the variance for the two groups.

Table 4.

Fit of Models Including Language Fundamentals and Non-Verbal Intelligence as Predictors of Working Memory Factors

Model	**Model χ^2**	**χ^2 Comparison of Previous and Current Models[1]**	**RMSEA and 90% CI**	**CFI**	***N*-adjusted BIC**
Noninvariant Slopes, Noninvariant Error (Co)variances	$\chi^2(179) = 219.71$, $p = .02$	--	.043 [.018, .061]	.93	13738.33
Noninvariant Slopes, Invariant Error (Co)variances	$\chi^2(185) = 225.24$, $p = .02$	$\chi^2(6) = 5.77$, $p = .45$	.042 [.017, .060]	.93	13730.84
Invariant Slopes, Invariant Error (Co)variances	$\chi^2(191) = 237.57$, $p = .01$	$\chi^2(6) = 10.24$, $p = .11$	.044 [.022, .061]	.92	13728.40

[1] This χ^2 difference test was conducted taking into account that the model χ^2 were based on maximum likelihood robust estimation (Sattora & Bentler, 2010).