Contemporary School Psychology (2020) 24:68–79
https://doi.org/10.1007/s40688-019-00246-z

Operationalizing the Simple View of Writing with the Wechsler Individual Achievement Test, 3rd Edition

Jason R. Parkin[1] · Craig L. Frisby[2] · Ze Wang[2]

Published online: 22 May 2019
© California Association of School Psychologists 2019

Abstract
The simple view of writing suggests that written composition results from oral language, transcription (e.g., spelling/handwriting), and self-regulation skills, coordinated within working memory. The model provides a number of implications for the interpretation of psychoeducational achievement batteries. For instance, it hypothesizes that writing skills are only partially related to each other through a hierarchy of levels of language (e.g., subword, word, sentence, discourse levels) and that transcription skills such as spelling mediate the effects of language skills on composition. We evaluated implications of the simple view of writing in the Wechsler Individual Achievement Test, 3rd Edition (WIAT-III). Using structural equation modeling, we established that WIAT-III writing tasks are only partially related to each other within both the battery's normative sample and an independent sample of students referred for special education. We also described how lower level writing skills mediated the effects of language skills on higher level writing skills. However, these effects varied across normative and referral samples.

Keywords Writing · WIAT-III · Psychoeducational assessment · Academic assessment · Invariance

The Wechsler Individual Achievement Test, 3rd Edition (WIAT-III; Wechsler 2009) is a popular, individually administered, standardized measure of academic achievement designed to assess the reading, writing, math, and oral language skills of individuals between ages 4 to 50 (Burns 2010). Despite publication almost a decade ago, little structural psychometric research has investigated the WIAT-III other than inclusion of its subtests as outcome variables explained by cognitive skills (e.g., Beaujean et al. 2014; Caemmerer et al. 2018). Indeed, detailed investigations of achievement batteries have lagged considerably compared to analyses of instruments that measure general cognitive ability (Dombrowski 2015). This research gap could be due to clinicians' tendencies to focus on latent factors when considering general cognitive functioning, while focusing on observed variables when evaluating academic achievement. To this point, initial WIAT-III construct validation efforts focused on correlations within WIAT-III achievement areas, with other published batteries, while less attention was devoted to the identification of latent factors within the battery (Breaux 2010). In addition, standardized test development and interpretation is often criticized as disconnected from psychological theory (Beaujean and Benson 2018). There are well-researched theories of academic skill development that can inform both the development and interpretation of academic batteries. An example of such a theory is "the simple view of writing" (Berninger et al. 2002). Our purpose is to describe the simple view of writing, its implications for psychoeducational assessment construct validity, and to apply those implications to evaluate aspects of WIAT-III test validity.

The Simple View of Writing

Initially, the simple view of writing conceptualized composition as the outcome of two skill areas, spelling and idea generation (Juel 1988). Relying on the Hayes and Flower (1980) model of adult writing, researchers expanded and clarified the model to describe the developmental trajectory of writing skill (Berninger 2009; Berninger and Amtmann 2003; Berninger and Swanson 1994; Hayes and Berninger 2009). The model includes three broad skill areas, transcription, text generation,

Standardization data from the *Wechsler Individual Achievement Test, Third Edition (WIAT-III)*, Copyright © 2009 NCS Pearson, Inc. Used with permission. All rights reserved.

✉ Jason R. Parkin
parkinj@seattleu.edu

[1] Seattle University, Seattle, WA, USA

[2] University of Missouri, Columbia, MO, USA

and self-regulation, collectively coordinated within working memory (Berninger and Winn 2006; Kim et al. 2015a). The effects of transcription, text generation, and self-regulation skills on students' composition are central to the model.

Transcription represents skills necessary to convert language (either heard from an external auditory source, or generated internally in one's mind) into print, specifically, spelling and handwriting (Hayes and Berninger 2009). Both handwriting and spelling intervention effects on writing have been investigated by comprehensive meta-analyses (Graham and Santangelo 2014; Santangelo and Graham 2016). Graham and Santangelo's handwriting analyses aggregated intervention effect sizes (ES) for students in Kindergarten through the ninth grade, stressing improvement in composition quality (ES = .84), length (ES = 1.33), and fluency (ES = .48). Handwriting was effectively supported via individualized instruction (ES = .69) and technology (such as using a tablet to copy letters; ES = .85). Their spelling analyses demonstrated that explicit instruction transferred to correct spelling within composition (ES = .94), though spelling instruction did not appear to increase general writing performance (ES = .19). An increase in general writing may also require explicit instruction pertaining to composition (Berninger et al. 2002).

Text generation represents a writer's skill constructing ideas, translating those ideas into words, sentences, or paragraphs, and includes topic and genre knowledge (Berninger et al. 2002). Text generation is often operationalized as oral language skills and represents students' collective knowledge of vocabulary, grammar, morphology, syntax, and language fluency (Kim et al. 2015b; Kim and Schatschneider 2017; McCutchen 2011). Researchers have demonstrated that language competence relates to effective writing across school grades (Berninger and Abbott 2010; Dockrell and Connelly 2009; Kim et al. 2011, 2015a, 2015b). Research investigating language intervention effects on writing appears significantly less developed than spelling/handwriting intervention (Shanahan 2006). Comparisons of students with specific language impairments to students with decoding impairments and typically developing peers suggest that oral language challenges may lead to shorter composition and higher rates of spelling and grammatical errors (Connelly et al. 2012; Dockrell et al. 2009; Puranik et al. 2007). Some investigators suggest that writing deficits can persist even after the remediation of language difficulties (Dockrell et al. 2009; Nauclér and Magnusson 2002).

Self-regulation skills such as goal-setting, self-assessment, and writing strategy instruction can support text planning, generation, and revision efforts (Santangelo et al. 2008). Indeed, explicit intervention to support self-regulation can increase the quality of writing, particularly when added to writing strategy instruction (ES = .59; Graham et al. 2012). Ritchey et al. (2016) stressed that the types of supports necessary to scaffold students' self-regulation when writing might vary based on the type of writing task. For instance, when spelling, the writer might require help sounding out a word or reviewing their spelling choices. Alternatively, when composing, writers might require support with graphic organizers, or other prompts for topics/subtopics.

It is important to stress that these skills are not necessarily independent of each other (Hayes and Berninger 2009). Transcription skills may constrain text generation (Graham et al. 1997; Hayes and Berninger 2009). For instance, when young writers dictate their thoughts, removing the need for transcription skills, they produce stronger text (De La Paz & Graham, 1995). When fluent writers transcribe in a novel, unpracticed way, their sentences become shorter and less sophisticated (Grabowski 2010). Berninger (1999) suggested that handwriting and spelling skills place more cognitive load on working memory in developing writers, limiting students' ability to translate ideas into language.

The concept of levels of language is also central to the simple view of writing. As language is not a single construct, listening comprehension, oral expression, reading, and writing reflect both interconnected and independent developmental systems (Berninger 2000). Berninger and Abbott (2010) demonstrated that these four language systems contain both shared and unique variance longitudinally and stressed that individual strengths and weaknesses in these language systems can be common (even in "typically developing" youth) and relatively stable over time. These language systems can be compared and contrasted not only expressively/receptively but also at subword, word, sentence, and text/discourse levels of languages. These levels are partially hierarchical; students can engage in higher levels of language without proficiency in lower levels, even though the units of higher levels are comprised of lower levels of language (Abbott et al. 2010; Berninger et al. 1988, 1994). As a case in point, Berninger et al. (1994) reported non-significant relationships between measures of spelling, sentence, and paragraph construction completed by intermediate grade students. The implication is that students may demonstrate intraindividual differences in performance across various levels of language, and performance at one level should only partially explain performance at another. The WIAT-III technical manual authors stressed this aspect of writing development, and provided subtest correlations, but did not investigate latent factors or theory-based structural relationships (Breaux 2010).

Implications for Validity of a Writing Battery

The WIAT-III includes subtests assessing writing performance at multiple language levels (Breaux 2010). Alphabet writing fluency provides a measure of fluent letter recall and legibility, writing at the subword level. Breaux and Lichtenberger (2016) stressed that it is not a handwriting task, though others have

categorized it as such (Drefs et al. 2013). It is certainly a measure of transcription skills. The measure can be administered to students up to grade 3. Spelling requires examinees to spell single words from dictation (though the earliest items include single letters) and reflects transcriptive writing at the word level for most examinees. It requires knowledge of letter/sound correspondence, prefixes, suffixes, and also homophones, and thus, effective performance will also tap subword language skills, semantics, and morphology. The sentence composition task is comprised of two components. Sentence combining requires examinees to convert short sentences into one sentence of greater complexity. Sentence building requires examinees to generate a sentence based on a target word. Collectively, performance on these sentence composition tasks is most likely influenced by an awareness of semantics, syntax, grammar, capitalization and punctuation, and spelling. At the text level, essay composition requires examinees to write an essay within 10 min. It provides scores representing theme and organization, word count, an aggregation of theme and word count, and a score reflecting grammar and mechanics.

As some researchers suggest that oral language can represent text generation skills (Kim and Schatschneider 2017), we included the WIAT-III oral language measures in these analyses. Within the battery, receptive language includes two components: one requiring the recognition of a picture definition of a word (receptive vocabulary) and the other requiring answers to questions about brief audio passages (oral discourse comprehension). Expressive language tasks include an expressive vocabulary measure, which requires examinees to say a word labeling a picture and its corresponding definition; a word fluency task, requiring examinees to quickly provide names of category exemplars; and a sentence repetition task, in which examinees restate sentences read by the examinee.

The simple view of writing provides a strong theoretical basis for test developers to evaluate aspects of construct validity and for clinicians to describe both the writing performance of struggling students as well as develop intervention strategies. As most achievement batteries include tests operationalizing the model's core constructs, the simple view of writing provides predictions that can be empirically evaluated via the relationships between subtests. First, the battery's oral language tasks should demonstrate effects on each written level of language (e.g., spelling, sentence construction, and composition). Second, as transcription skills may constrain students' text generation, measures of transcription should mediate effects of language on higher levels of language. Third, as language levels are only partially hierarchical, the effects of one language level on another should only be low to moderate. Lower levels of language should demonstrate both direct and indirect effects on higher language levels.

Our purpose is to evaluate the extent to which the WIAT-III tasks operationalize the aforementioned hypotheses. We used the battery's normative sample and an independent sample of students referred for special education to formally test invariance between the two samples. Invariance testing can support whether language/writing constructs established in the normative sample represent the same skills or abilities in students referred for testing due to academic difficulties (Wicherts 2016). Results should add to the construct validity base of the WIAT-III.

Method

Participants

WIAT-III Standardization Sample The WIAT-III standardization sample was stratified to approximate the US population in 2005, as reflected in the U.S. Bureau of the Census on the basis of grade, age, sex, race/ethnicity, parent education level, and geographic region (Breaux 2010). Because students vary in the writing measures they complete based on their grade level, we analyzed the performance of students in students from grades 1–3 ($n = 668$) in one model and grades 3–12 ($n = 2226$) in a second model. These grade ranges conform to differences in the WIAT-III subtest administration procedures.

Referral Sample We gathered a sample of students from a suburban school district in the Pacific Northwest who completed all WIAT-III oral language and writing subtests through special education evaluation. Other portions of this sample were reported by Parkin (2018). All participants completed the battery during the same school year. Specifically, 143 students completed subtests for grades 1–3, and 345 students completed subtests for grades 3–12. Most of the same third grade students were included in both age groups, though some were excluded from one group or the other because they did not complete either the alphabet writing fluency task or the essay composition task. The grades 1–3 group included two third grade students not in the grades 3–12 group. The grades 3–12 group included four third grade students not in the grades 1–3 group. Socioeconomic status data is not available for the referral sample. Approximately 12% of students who attend the school district qualified for free or reduced Lunch, according to district information. We compare available demographic data for these samples in Table 1.

Measures

We used oral language and writing measures from the WIAT-III to operationalize text generation, transcription, and composition skills. We provide a description of these tasks and their reliability coefficients in Table 2.

Table 1 Demographic characteristics of WIAT-III normative and referral samples

Variable	Referral (grades 3–12)		Referral (grades 1–3)		Norm	
	n	%	*n*	%	n	%
Sex						
Male	212	62	92	64	1576	71
Female	132	38	51	36	1544	69
Grade						
1			48	34	226	10
2			52	36	225	10
3	44	13	43	30	225	10
4	75	22			223	10
5	56	16			222	10
6	37	11			223	10
7	33	10			229	10
8	37	11			217	10
9	30	9			221	10
10	22	6			227	10
11	9	3			219	10
12	1	0			222	10
Race/ethnicity						
American Indian	2	1	0	0		
Asian	41	12	13	9	108	5
Black	6	2	3	2	459	21
Hispanic	71	21	33	23	576	26
Multiracial	23	7	7	5		
White	201	58	87	61	1872	84
Other					105	5
Language						
English	258	75	105	73		
Spanish	44	13	21	15		
Other	42	12	17	12		
Evaluation type						
Initial	149	43	81	57		
Reevaluation	195	57	62	43		
Evaluation outcome						
504	18	5	4	3		
Autism	35	10	11	8		
Communication disorder	8	2	8	6		
Developmental delay	0		13	9		
No eligibility	27	8	13	9		
Emotionally disabled	9	3	1	1		
Health impairment	109	32	35	24		
Intellectual disability	4	1	2	1		
Multihandicapped	1	0	0	0		
SLD	133	39	56	39		
Areas of specially designed instruction						
Basic reading	96	28	57	40		
Reading comp.	131	38	49	34		
Reading fluency	96	28	59	20		
Listening comp.	55	16	13	9		
Oral expression	43	13	16	11		
Written expression	173	50	69	48		
Math calculation	140	41	47	33		
Math prob. solving	165	48	52	36		
Soc/emo/beh	112	33	52	36		
Organization	142	41	18	13		
Adaptive skills	24	7	7	5		
Physical therapy	4	1	3	2		
Occupational	27	8	40	28		
Speech/language	77	22	51	36		

Procedure

We obtained permission from the publisher to analyze the WIAT-III standardization sample. We gathered archived WIAT-III test data on referred students attending a large school district in the Pacific Northwest. Forty-seven certified special education teachers trained in the administration of the WIAT-III gathered data from students as part of a special

Table 2 WIAT-III subtest-component descriptions

Task name	Description	Reliability
Essay composition	Score is the total rubric points across introduction, conclusion, paragraphs, transitions, "reasons why", elaborations, and a total word count.	.88[b]
Sentence composition	Examinee combines simple sentences into a complex sentence and generates a sentence from a target word. Both sentence types are scored based on their mechanics (including spelling) and semantics/grammar.	.87[b]
Spelling	Examinee writes a word after hearing it in a sentence.	.95[a]
Alphabet writing fluency	Examinee writes upper or lowercase letters in a short time limit.	.69[b]
Expressive vocabulary	Examinee views a picture on a page and must verbalize the word shown on the page when prompted by the examiner.	.73[a]
Oral word fluency	The examiner asks the examinee to say as many words as s/he can within a certain category in 1 min.	.75[b]
Sentence repetition	The examinee repeats verbatim a sentence read by the examiner.	.83[a]
Oral discourse comprehension	The examinee answers factual and inferential questions based on a passage they just heard.	.82[a]
Receptive vocabulary	The examinee points to a picture that best represents a word they just heard.	.69[a]

[a] Split-half

[b] Test/retest

education eligibility evaluation process during the 2015–2016 school year. Teachers administered between one and 35 assessments with a mean of 10.1. Order of subtest administration is not known. The battery was administered during the standard school day. The specific number of sessions for each student is not known. We gathered demographic data from a district database and test scores from the publisher's online scoring system.

Analyses

We conducted all analyses with Mplus Version 8 (Muthén and Muthén 1998-2017). Within each age group, our analyses proceeded in three stages.

Measurement Invariance We evaluated configural, metric, and scalar invariance across normative and referral samples. Because the WIAT-III provides single measures of subword, word/spelling, and discourse/text skills, we modeled each of these factors with a single indicator. To ensure model identity, we constrained disturbance variance in these factors to 0.

Structural Equation Modeling Next, for each age group, we modeled the effects of the oral language latent factor on each level of language factor. We also evaluated the effects of lower language levels on higher levels. Figure 1 depicts the model for grades 1–3, and Fig. 2 describes the model for grades 3–12.

Consistency of Effects Across Samples We used Wald's test to compare the equivalency of effects across each sample. This included five coefficients in the grades 1–3 model and six coefficients in the grades 3–12 model.

Model Fit For model fit statistics, we relied on multiple measures (Keith 2015). The comparative fit index (CFI) and the Tucker–Lewis index (TLI) indicate a strong fit when values reach .95 or higher. The root mean square error of approximation (RMSEA) when values are lower than .06 and standardized root mean square residual (SRMR) when values are lower than .08 reflect appropriate fit (Hu and Bentler 1999). Changes in model fit statistics have been recommended for testing measurement invariance (Chen 2007; Cheung and Rensvold 2002). These recommendations include a decrease in CFI by at least .01, supplemented by an increase in RMSEA by at least .015 for testing factor loading invariance or for testing intercept or residual invariance, if a less constrained model is to be selected. When evaluating model effects, Keith (2015) suggests that effects between .05 and .10 are small, effects between .10 and .25 are moderate, and effects larger than .25 are large.

Results

Descriptive Statistics

We report descriptive statistics of manifest variables in Table 3. As would be expected, participants in the referral sample demonstrated lower performance on almost all variables. The referral sample also displayed greater variability in scores.

Measurement Invariance

Measurement invariance results for the grades 1–3 group are in Table 4. Using normative and referral sample type as the grouping variable, the configural, metric, and scalar invariance

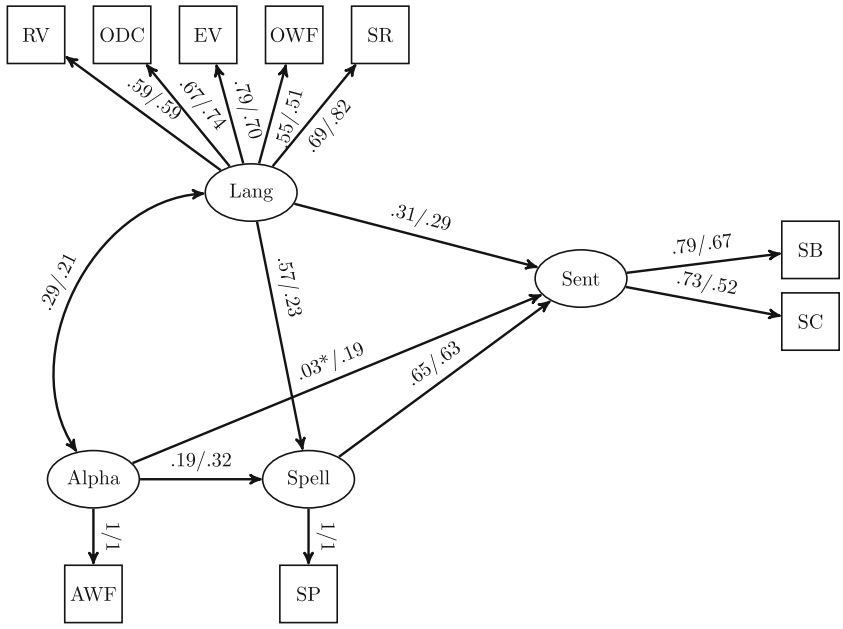

Fig. 1 Standardized effects comparing samples on grades 1–3 WIAT-III language and writing measures. Coefficients represent the normative/referral samples. Asterisk indicate statistically non-significant coefficient. Disturbance factors not included to easy readability. RV receptive vocabulary, ODC oral discourse comprehension, EV expressive vocabulary, OWF oral word fluency, SR sentence repetition, AWF alphabet writing fluency, SP spelling, SB sentence building, SC sentence combining, Lang language, Sent sentence writing, Alpha alphabet writing, Spell spelling

models fit the data well. The ΔCFI from the configural invariance model to the metric invariance model was − .001, and the ΔCFI from the metric invariance model to the scalar invariance model was − .023. The ΔRMSEA from the configural invariance model to the metric invariance model was 0, and ΔRMSEA from the metric invariance model to the scalar invariance model was .019. These results suggest that between configural, metric, and scalar invariance models, metric invariance model should be selected. Further examining the scalar invariance model indicated that two intercept constraints for the sentence combining and oral word fluency subtest components should be released. After releasing the two constraints, the partial scalar invariance model fit the data well (CFI = .976, TLI = .968, SRMR = .058, RMSEA = .051 with 90% CI [.037, .064]) and not worse than the metric invariance model. We retained the partial invariance model for subsequent analyses.

Measurement invariance results for the grades 3–12 group are in Table 5. The configural and metric invariance models fit the data well, and the scalar invariance model fits the data acceptably. The ΔCFI from the configural invariance model to the metric invariance model was .001, and the ΔCFI from the metric invariance model to the scalar invariance model was − .032. The ΔRMSEA from the configural invariance model to the metric invariance model was − .003, and the ΔRMSEA from the metric invariance model to the scalar invariance model was .028. These values suggest that between

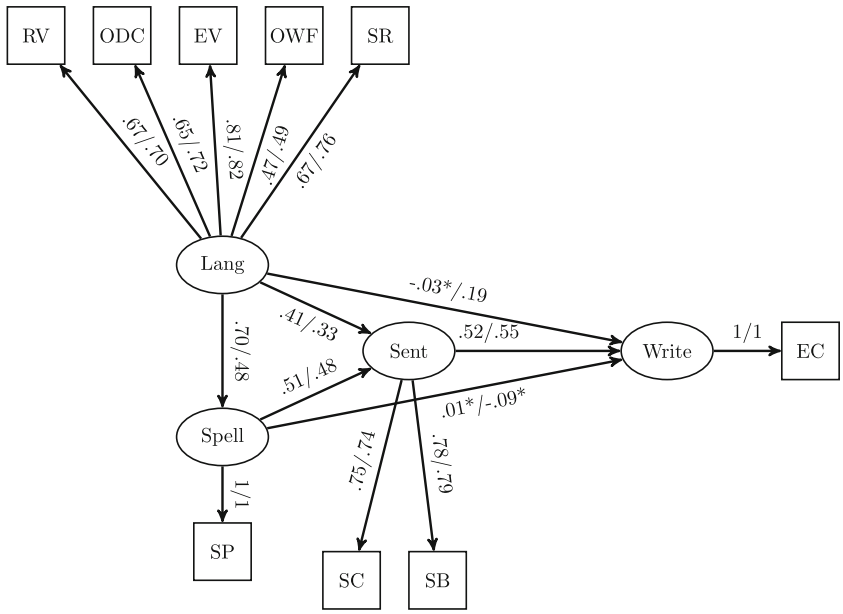

Fig. 2 Standardized effects comparing samples on grades 3–12 WIAT-III language and writing measures. Coefficients represent the normative/referral samples. Asterisk indicates statistically non-significant coefficient. Disturbance factors not included to easy readability. RV receptive vocabulary, ODC oral discourse comprehension, EV expressive vocabulary, OWF oral word fluency, SR sentence repetition, EC essay composition, SP spelling, SB sentence building, SC sentence combining, Lang language, Sent sentence writing, Spell spelling, Write essay writing

Table 3 Descriptive statistics across WIAT-III normative and referral samples

	Grade 1–3								Grade 3–12							
	Mean		SD		Skewness		Kurtosis		Mean		SD		Skewness		Kurtosis	
Subtest	Norm	Ref	Norm	Ref	Norm	Ref	Norm	Ref	Norm	Ref	Norm	Ref	Norm	Ref	Norm	Ref
RV	99.77	97.97	15.33	17.60	−.13	−.27	.15	− .50	100.16	97.38	15.12	17.04	−.05	−.46	.15	.36
ODC	100.92	95.44	14.81	14.03	.19	−.18	.95	1.04	100.46	90.34	15.32	15.61	.04	−.26	−.16	.58
EV	99.48	91.20	14.94	18.99	−.21	.24	.26	.49	100.08	89.15	15.06	16.71	−.15	−.07	−.01	−.04
OWF	99.55	101.41	15.16	17.43	.00	.42	1.07	.49	100.05	101.98	15.36	17.35	.06	.18	.53	.52
SR	99.93	92.02	15.36	14.30	.08	.13	.61	.07	100.31	89.52	15.44	15.49	.11	−.19	.87	.63
AWF	99.96	98.81	15.70	16.81	.08	.26	.48	−.27								
SP	100.59	91.22	15.24	15.03	−.27	.53	.74	1.97	99.68	89.72	15.65	14.74	−.17	−.07	−.01	.24
SC	97.92	93.08	15.84	15.83	−.15	.26	−.22	−.01	99.88	97.16	15.14	16.76	−.30	.04	−.15	−.34
SB	99.152	86.33	15.39	13.29	−.06	.50	.27	.49	100.05	86.93	15.16	16.89	−.50	.02	.21	−.34
EC									99.79	96.23	15.36	17.11	−.12	−.38	.29	1.29

RV receptive vocabulary, *ODC* oral discourse comprehension, *EV* expressive vocabulary, *OWF* oral word fluency, *SR* sentence repetition, *AWF* alphabet writing fluency, *SP* spelling, *SC* sentence combining, *SB* sentence building, *EC* essay composition

configural, metric, and scalar invariance models, metric invariance model should be selected. Further examining the scalar invariance model indicated that three intercept constraints for sentence combining, oral word fluency, and receptive vocabulary subtest components should be released. After releasing the three constraints, the partial scalar invariance model fits the data well (CFI = .980, TLI = .974, SRMR = .043, RMSEA = .049 with 90% CI [.042, .056]) and not worse than the metric invariance model. Therefore, we retained the partial invariance model for subsequent analyses.

From the final partial scalar invariance models, we obtained latent mean differences between the normal sample and the referral sample. For the grades 1–3 group, the referral group displayed statistically significantly lower language, spelling, and sentence writing factor scores, compared to the normative sample; the norm sample and the referral sample were comparable on alphabet writing factor. For the grades 3–12 group, the referral group displayed statistically significantly lower language, spelling, sentence writing, and essay composition factor scores, compared to the normative sample.

Structural Equation Models

Based on the final measurement model for each age group, we evaluated path effects from the language factor to the writing factors and between levels of language. We also investigated differences between effects across samples via Wald's test: five structural paths in the grades 1–3 sample, $\chi^2(5) = 28.52$, $p < .001$ and six structural paths in the grades 3–12 sample, $\chi^2(6) = 53.77$, $p < .001$.

For each age group, the model fit for the two group structural regression model was the same as the model fit for the partial scalar invariance model. Figure 1 depicts the path model for the grades 1–3 sample, illustrating the effects of oral language on levels of language and the interrelationships between writing tasks. The language factor demonstrated a moderate direct effect on sentence writing. Its effect on spelling varied by sample, moderate in the norm sample, and smaller in the referral sample. The language factor also displayed a small to moderate relationship level with the alphabet factor. Regarding levels of language, the subword alphabet writing factor demonstrated a small to moderate effect on the word

Table 4 Model fit for WIAT-III measurement invariance models for grades 1–3 group

	# Para.	Chi-square	df	CFI	TLI	RMSEA	90% CI		SRMR
Configural invariance	60	93.474	48	.981	.971	.048	.034	.063	.051
Metric invariance	56	100.307	52	.980	.972	.048	.034	.062	.061
Scalar invariance	51	159.851	57	.957	.945	.067	.055	.079	.082
Partial scalar invariance	53	112.905	55	.976	.968	.051	.037	.064	.058

All chi-square values were statistically significant at the .001 level

CFI comparative fit index, *TLI* Tucker–Lewis index, *RMSEA* root mean square error of approximation, *SRMR* Standardized Root Mean Square Residual

Table 5 Model fit for WIAT-III measurement invariance models for grades 3–12 group

	# Para.	Chi-square	df	CFI	TLI	RMSEA	90% CI		SRMR
Configural invariance	60	211.891	48	.980	.971	.052	.045	.059	.040
Metric invariance	56	215.623	52	.981	.973	.049	.043	.056	.043
Scalar invariance	51	488.983	57	.949	.935	.077	.071	.083	.072
Partial scalar invariance	54	220.544	54	.980	.974	.049	.042	.056	.043

All chi-square values were statistically significant at the .001 level

CFI comparative fit index, *TLI* Tucker–Lewis index, *RMSEA* root mean square error of approximation, *SRMR* Standardized Root Mean Square Residual

level spelling factor. It displayed a negligible to small effect on the sentence level factor for the two samples. The spelling factor demonstrated a large effect on the sentence factor. Collectively, the model explained 42% of variance in the spelling factor for the normative group, but only 19% in the referral sample. For sentence writing, the model explained 78% of variance in the normative sample and 74% of variance in the referral sample.

Figure 2 displays path effects for the grades 3–12 sample. Sample type moderated the effect of language on spelling. In the normative sample, the effect was large, while in the referral sample, the effect was significantly smaller. The model explained 49% of spelling variance in the normative sample and 23% of variance in the referral sample. Language displayed a moderate effect on sentence writing and a negligible to small effect on essay composition, again moderated by sample. The language-to-composition effect was not significant in the normative sample and small in the referral sample. In terms of levels of language, the word level spelling factor displayed a moderate effect on the sentence writing factor and no direct effect on the essay composition factor. Its effect was completely mediated by the sentence factor. The sentence factor demonstrated a moderate effect on the essay composition factor. The model explained 49% of spelling variance in the normative sample and 23% in the referral sample. Regarding sentence writing, the model explained 72% of variance in the normative sample and 48% of variance in the referral sample. It explained 25% of essay composition variance in the normative sample and 38% in the referral sample.

Discussion

The simple view of writing suggests that text generation skills, operationalized as oral language, should demonstrate effects on word, sentence, and composition skills. The simple view of writing also suggests that transcription skills (e.g., spelling/handwriting) should mediate those effects. The model conceptualizes writing tasks via hierarchical levels of language. Subword, word, and sentence skills may only partially influence performance on higher levels in the hierarchy. We investigated these implications of the simple view of writing within the WIAT-III, using latent factors to establish both the effects of a language factor on written expression tasks and the effects between writing tasks in the battery. We also compared these implications across the battery's normative sample and an independent sample of students referred for special education.

Our measurement model established that WIAT-III writing tasks can indeed be interpreted in a manner that is consistent with a level of language model (Berninger et al. 1994). In the younger sample, a model specifying subword, word, and sentence writing skills provided a strong fit to the model. Similarly, in grades 3–12, a model with word, sentence, and composition level tasks reflected a strong model fit. These results are generally consistent with longitudinal analyses conducted with the second edition of the WIAT from a levels of language perspective (Abbott et al. 2010; Wechsler 2001). Importantly, the measurement model demonstrated metric and partial scalar invariance across normative and referral samples. These findings indicate that the latent factors may be interpreted with the same meaning across samples. However, mean score differences across groups in the sentence combining and oral word fluency measures (and in the grade 3–12 groups, the receptive vocabulary measure) may be more related to those tasks, rather than the latent factor they represent.

Though the measurement model was consistent across samples, structural equation modeling indicated that the effects of latent language and writing variables varied across samples. We provide numerical comparison of the effects of language skills and lower level writing skills on writing performance in Table 6. In both samples and across both models, language skills demonstrated direct effects on word and sentence level skills and an association with subword performance in the younger grade model. Language effects on spelling varied by sample. In the normative sample, language and spelling appear to develop closely together, but they are less associated with each other in the referral sample. These results likely underscore a core phonological deficit in youth struggling to develop word-level literacy skills (Fletcher et al. 2007). Language and spelling affected sentence writing in the same way across samples and within models. Collectively, these results indicate that

Table 6 Comparison of standardized effects across WIAT-III normative/referral samples

	Spell			Sent			Write		
	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total
Grades 1–3									
LANG	.57/.23		.57/.23	.31/.29	.37/.15	.68/.43			
ALPHA	.19/.32		.19/.32	.03/.19	.12/.20	.15/.40			
SPELL				.65/.63		.65/.63			
R^2			.42/.19			.78/.74			
Grades 3–12									
LANG	.70/.48		.70/.48	.41/.33	.36/.23	.77/.56	.03/.19	.41/.26	.37/.45
SPELL				.51/.48			.01/− .09	.26/.26	.28/.17
SENT							.52/.55		.52/.55
R^2			.49/.23			.72/.48			.25/.34

transcription skills represent a significant bottleneck on essay writing, perhaps increasing the difficulty of language production (Berninger 1999; Bourdin and Fayol 1994; Connelly et al. 2005). The bottleneck appeared more extreme in the normative sample due to a smaller association between spelling and language in the referral sample. These results are consistent with our hypotheses stemming from the simple view of writing's description of the text generation/transcription relationship. Lower level writing skills largely mediated language effects on higher order skills. At the same time, the larger, direct effect for language on composition in the referral sample could suggest a compensatory process. Perhaps, writers with stronger language skills can more easily select words and phrases they can spell, reducing the transcription bottleneck.

While WIAT-III writing tasks demonstrated partial independence, these measures appear more closely associated than those described by others (Berninger et al. 1994). Berninger et al. (1994) used writing tasks similar to those contained in the WIAT-III but reported non-significant relationships between them. One reason may be differences in scoring. For instance, here the spelling/sentence writing relationship may be higher than Berninger's investigation, because some types of spelling errors are included as part of the scoring criteria in sentence composition (Breaux 2010). In contrast, essay composition scoring includes a word count and a rubric to evaluate organization and theme development; neither component contains criteria explicitly associated with lower language levels. This could explain why the model explained the less variance in the essay composition variable than in the spelling or sentence writing variables.

Consistency with Previous Research

These results are consistent with a number of prior investigations. Kim et al. (2015b) noted significant effects for spelling on a latent writing quality variable in a small sample of Korean students. Their oral language factor, a discourse-level measure more narrow than the factor we used here, only approached significance, likely due to the small sample size. However, they did not model a mediation effect for spelling, nevertheless noting an association between spelling and discourse language. Using the WIAT-II, Berninger and Abbott (2010) reported effects for listening comprehension and oral expression skills on written expression in multiple grade levels, though this study did not include a measure of spelling skills. Other researchers have investigated effects of cognitive performance on writing skills (Caemmerer et al. 2018; Cormier et al. 2016; Hajovsky et al. 2018). If crystallized intelligence (Gc) might be construed as oral language skills, Hajovsky et al. (2018) demonstrated language effects on writing moderated by grade level in the Kaufman Test of Educational Achievement, Second Edition (KTEA-2; Kaufman and Kaufman 2004). In comparison, using Wechsler batteries, Caemmerer et al. (2018) noted effects for Gc on spelling, but not for essay writing, when included in analyses with other cognitive variables, and Cormier et al. (2016) describe similar differences in Gc effects across basic writing and written expression. These analyses regressed single writing tasks on cognitive variables and may not account for the effects of academic tasks on each other, a key implication from the simple view of writing.

Implications for Practice and Test Interpretation

These analyses suggest partial independence between levels of written language, as assessed by the WIAT-III. Clinicians should expect variability in these skills when evaluating examinee writing performance. To describe that performance, clinicians may need to focus interpretation on specific tasks. At the same time, we stress that subtests may not provide the level of reliability necessary to make high-stake decisions; in this battery, only spelling demonstrates a reliability level of .90 or higher (Breaux 2010).

Levels of language interpretation of writing tasks may have implications for the interpretation of a writing composite. Schneider (2013) described two ways to consider the relationships between subtests and composites. Generality reflects the idea that some abilities influence a wider range of functioning. For example, pertaining to cognitive measures, "g" would influence most all cognitive tasks, a broad ability would influence tasks reflective of a general class of abilities (e.g., only visual/spatial skills), and a narrow ability would influence very specific tasks (e.g., Spatial Relations). In comparison, abstractness represents the degree of complexity in a task. A more abstract task may require the coordination of multiple skills/abilities. We contend that the levels of language fit described here suggests that writing skills might effectively be conceptualized in this way. Spelling may represent the coordination of letter retrieval, phonological, orthographic, and morphological skills, as well as semantic/vocabulary knowledge (Berninger et al. 2006), while sentence writing adds aspects of grammar and syntax. Composition may require the coordination of even more skills.

Limitations and Future Directions

There are a number of limitations associated with these analyses. First, we recognize that the tested models reflect an incomplete operationalization of the simple view of writing. They omit indicators of handwriting, working memory, and self-regulation skills. Like spelling, handwriting represents a transcription skill and likely mediates the relationship between oral language and levels of written expression. However, working memory and self-regulation, skills that coordinate other skills, may be more challenging to conceptualize. Do these skills explain independent variance in writing, or do they moderate or mediate the influence of other skills? Kim and Schatschneider (2017) reported that language skills mediate working memory effects on composition. Poch and Lembke (2017) provided a latent variable analysis including self-regulation and working memory skills alongside writing, though noted a poor model fit. In an exploratory analysis, they reported that working memory and lower-order writing skills loaded together and self-regulation, planning, handwriting, and essay composition loaded on a second factor. However, the authors noted that their results may have been impacted by a small sample size for the analyses they employed. It may also be necessary to expand these models beyond variables included in the simple view of writing. Levels of language conceptualization of writing suggest that there may be unique predictors of performance at each language level. For instance, phonological skills may explain variance in spelling, but not in essay writing, controlling for spelling. These models could be expanded to investigate such predictors, particularly at the essay level; language and lower-level writing skills only explained 25% of essay writing variance in the normative sample.

Second, the referral samples we used in these analyses come from only one district and may not generalize to others. Although the samples demonstrated a measurement model consistent with the normative sample, that generalizability may be due to the inclusion of referred students without writing difficulties. The referral samples were heterogeneous in that they included students with numerous types of disabilities and educational needs. Only approximately 50% of the referral samples included students that required specially designed instruction in writing. These results may differ if only students with writing challenges were analyzed.

Third, we evaluated the consistency of a specific model across samples. However, there may be other models that reflect an appropriate model fit. Finally, there may be developmental effects that require further investigation. Berninger (1999) highlighted that effects of transcription skills on composition decrease across development and Hajovsky et al. (2018) highlighted grade level moderation of Gc-to-writing effects.

Fourth, it is important to stress that though we highlighted mediation effects, the data here are ultimately correlational, which limits our ability to make causal inferences. However, we reported both intervention research and longitudinal analyses that support causal relationships between these correlations (Abbott et al. 2010; Graham et al. 2012; Graham and Santangelo 2014; Santangelo and Graham 2016).

Conclusion

The WIAT-III appears to operationalize important aspects of the simple view of writing effectively. Its writing measures appear to reflect multiple levels of language, though it is possible the spelling/sentence writing relationship is exacerbated, because spelling is an explicit scoring criterion within sentence writing. Spelling and sentence writing tasks mediated effects of oral language on composition, as predicted by the simple view of writing, though mediation occurred to a lesser degree in the referral sample. Collectively, clinicians should consider the simple view of writing in their interpretation of the battery.

Acknowledgments The authors wish to thank William Schryver and the team at Pearson for allowing access to the WIAT-III standardization sample.

Compliance with Ethical Standards

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of Interest The authors declare that there is no conflict of interest.

References

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*, 281–298. https://doi.org/10.1037/a0019318.

Beaujean, A. A., & Benson, N. F. (2018). Theoretically-consistent cognitive ability test development and score interpretation. *Contemporary School Psychology, 23,* 126.137. https://doi.org/10.1007/s40688-018-0182-1

Beaujean, A. A., Parkin, J., & Parker, S. (2014). Comparing Cattell-Horn-Carroll factor models: differences between bifactor and higher order factor models in predicting language achievement. *Psychological Assessment, 26*(3), 789–805. https://doi.org/10.1037/a0036745.

Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: automatic and constructive processes. *Learning Disability Quarterly, 22*, 99–112.

Berninger, V. W. (2000). Development of language by hand and its connections with language by ear, mouth, and eye. *Topics in Language Disorders, 20*(4), 65–84. https://doi.org/10.1097/00011363-200020040-00007.

Berninger, V. W. (2009). Highlights of programmatic, interdisciplinary research on writing. *Learning Disabilities Research & Practice, 24*(2), 69–80. https://doi.org/10.1111/j.1540-5826.2009.00281.x.

Berninger, V. W., & Abbott, R. D. (2010). Listening comprehension, oral expression, reading comprehension, and written expression: related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology, 102*(3), 635–651. https://doi.org/10.1037/a0019319.

Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345–363). New York: Guilford.

Berninger, V., & Swanson, H. L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. In J.S. Carlson & E.C. Butterfield (Eds.), *Children's writing: toward a process theory of development of skilled writing* (pp. 57–81). Binsley, UK: Emerald Group Publishing Limited.

Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald's (Eds.), *Handbook of writing research* (pp. 96–114). New York: Guilford.

Berninger, V. W., Proctor, A., De Bruyn, I., & Smith, R. (1988). Berninger, Proctor, De Bruyn, & Smith, 1988 JSP.pdf. *Journal of School Psychology, 26*, 341–357.

Berninger, V. W., Mizokawa, D. T., Bragg, R., Cartwright, A., & Yates, C. (1994). Intraindividual differences in levels of written language. *Reading & Writing Quarterly, 10*, 259–275.

Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., Hawkins, J. M., & Graham, S. (2002). Teaching spelling and composition alone and together: implications for the simple view of writing. *Journal of Educational Psychology, 94*, 291–304. https://doi.org/10.1037/0022-0663.94.2.291.

Berninger, V.W., Abbott, R.D., Jones, J., Wolf, B.J., Gould, L., Anderson-Youngstrom, M., Shimada, S., & Apel, K. (2006) Early development of language by hand: Composing, reading, listening, and speaking connections; Three letter-writing modes; and fast mapping in spelling. *Developmental Neuropsychology, 29*, 61-92 Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology, 29*, 591–620.

Breaux, K. C. (2010). *WIAT-III technical manual with adult norms*. Bloomington: Pearson.

Breaux, K. C., & Lichtenberger, E. O. (2016). *Essentials of KTEA-3 and WIAT-III assessment*. Hoboken: John Wiley & Soncs, Inc.

Burns, T. G. (2010). Wechsler individual achievement test-III: what is the "gold standard" for measuring academic achievement? *Applied Neuropsychology, 17*, 234–236.

Caemmerer, J. M., Maddocks, D. L. S., Keith, T. Z., & Reynolds, M. R. (2018). Effects of cognitive abilities on child and youth academic achievement: evidence from the WISC-V and WIAT-III. *Intelligence, 68*, 6–20. https://doi.org/10.1016/j.intell.2018.02.005.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. https://doi.org/10.1080/10705510701301834.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Connelly, V., Dockrell, J. E., & Barnett, J. (2005). The slow handwriting of undergraduate students constrains overall performance in exam essays. *Educational Psychology, 25*, 97–105.

Connelly, V., Dockrell, J. E., Walter, K., & Critten, S. (2012). Predicting the quality of composition and written language bursts from oral language, spelling, and handwriting skills in children with and without specific language impairment. *Written Communication, 29*(3), 278–302. https://doi.org/10.1177/0741088312451109.

Cormier, D. C., Bulut, O., McGrew, K. S., & Frison, J. (2016). The role of Cattell-Horn-Carroll (CHC) cognitive abilities in predicting writing achievement during the school-age years. *Psychology in the Schools, 53*, 787–803. https://doi.org/10.1002/pits.21945.

De La Paz, S., & Graham, S. (1995). Dictation: Applications to writing for students with learning disabilities. *Advances in Learning and Behavioral Disorders, 9*, 227-247.

Dockrell, J.E., & Connelly, V. (2009). The impact of oral language skills on the production of written text. In *BJEP monograph series II, number 6-teaching and learning writing, 45*, 45–62. https://doi.org/10.1348/000709909X421919.

Dockrell, J. E., Lindsay, G., & Connelly, V. (2009). The impact of specific language impairment on adolescents' written text. *Exceptional Children, 75*(4), 427–446.

Dombrowski, S. C. (2015). Exploratory bifactor analysis of the WJ-III achievement at school age via the Schmid–Leiman orthogonalization procedure. *Canadian Journal of School Psychology, 30*(1), 34–50. https://doi.org/10.1177/0829573514560529.

Drefs, M. A., Beran, T., & Fior, M. (2013). Methods of assessing academic achievement. *The Oxford handbook of child psychological assessment.* https://doi.org/10.1093/oxfordhb/9780199796304.013.0023.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: from identification to intervention*. New York: Guilford Press.

Grabowski, J. (2010). Speaking, writing, and memory span in children: output modality affects cognitive performance. *International Journal of Psychology, 45*, 28–39. https://doi.org/10.1080/00207590902914051.

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703–1743. https://doi.org/10.1007/s11145-014-9517-0.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: a new methodological approach. *Journal of Educational Psychology, 89*, 170–182. https://doi.org/10.1037/0022-0663.89.1.170.

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879–896. https://doi.org/10.1037/a0029185.

Hajovsky, D. B., Villeneuve, E. F., Reynolds, M. R., Niileksela, C. R., Mason, B. A., & Shudak, N. J. (2018). Cognitive ability influences

on written expression: evidence for developmental and sex-based differences in school-age children. *Journal of School Psychology, 67*, 104–118. https://doi.org/10.1016/j.jsp.2017.09.001.

Hayes, J. R., & Berninger, V. W. (2009). Relationships between idea generation and transcription: how the act of writing shapes what children write. In *Traditions of writing research*. https://doi.org/10.4324/9780203892329.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organisation of the writing process. In *Cognitive processes in writing* (Vol. 2, pp. 258–260). https://doi.org/10.1136/bmj.2.2169.258.

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), 1–55.

Juel, C. (1988). Learning to read and write: a longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437–447. https://doi.org/10.1037/0022-0663.80.4.437.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman test of educational achievement* (Second ed. Ed (KTEA-II)). Circle Pines: American Guidance Service.

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

Kim, Y. S. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: a direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology, 109*, 35–50. https://doi.org/10.1037/edu0000129.

Kim, Y. S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: an exploratory study. *Learning and Individual Differences, 21*, 517–525. https://doi.org/10.1016/j.lindif.2011.06.004.

Kim, Y. S., Al Otaiba, S., & Wanzek, J. (2015a). Kindergarten predictors of third grade writing. *Learning and Individual Differences, 37*, 27–37. https://doi.org/10.1016/j.lindif.2014.11.009.

Kim, Y. S. G., Park, C., & Park, Y. (2015b). Dimensions of discourse level oral language skills and their relation to reading comprehension and written composition: an exploratory study. *Reading and Writing, 28*(5), 633–654. https://doi.org/10.1007/s11145-015-9542-7.

McCutchen, D. (2011). From novice to expert: implications of language skills and writing relevant knowledge for memory during the development of writing skill. *Journal of Writing Research, 3*, 51–68.

Muthén, L. K., & Muthén, B. O. (1998-2017). Mplus user's guide (8th ed.). Los Angeles: Muthén & Muthén.

Nauclér, K., & Magnusson, E. (2002). How do preschool language problems affect language abilities in adolescence? In F. Windsor, M.L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics*. (pp. 99–114). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Parkin, J. R. (2018). Wechsler individual achievement test–third edition oral language and reading measures effects on reading comprehension in a referred sample. *Journal of Psychoeducational Assessment, 36*(3), 203–218. https://doi.org/10.1177/0734282916677500.

Poch, A. L., & Lembke, E. S. (2017). A not-so-simple view of adolescent writing. *International Journal for Research in Learning Disabilities, 3*, 27–44. https://doi.org/10.28987/irjrld.3.2.27.

Puranik, C. S., Lombardino, L. J., & Altmann, L. J. (2007). Writing through retellings: an exploratory study of language-impaired and dyslexic populations. *Reading and Writing, 20*(3), 251–272. https://doi.org/10.1007/s11145-006-9030-1.

Ritchey, K. D., McCaster, K. L., Al Otaiba, S., Puranik, C. S., Kim, Y. S. G., Parker, D. C., & Ortiz, M. (2016). Indicators of fluent writing in beginning writers. In K. D. Cummings & Y. Petscher (Eds.), *The fluency construct* (pp. 21–61). New York: Springer. https://doi.org/10.1007/978-1-4939-2803-3_2.

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review, 28*, 225–265. https://doi.org/10.1007/s10648-015-9335-1.

Santangelo, T., Harris, K. R., & Graham, S. (2008). Using self-regulated strategy development to support students who have "trubol giting thans into werds". *Remedial and Special Education, 29*, 78–89. https://doi.org/10.1177/0741932507311636.

Schneider, W. J. (2013). Principles of assessment of aptitude and achievement. In L. McKee, D. J. Jones, R. Forehand, & J. Cuellar (Eds.), *The Oxford handbook of child psychological assessment* (pp. 286–330). New York: Oxford University Press.

Shanahan, T. (2006). Relations among oral language, reading, and writing development. In C. MacAuthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 171–183). New York: Guilford.

Wechsler, D. (2001). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio: Psychological Corp.

Wechsler, D. (2009). *Wechsler Individual Achievement Test* (3rd ed.). San Antonio: NCS Pearson, Inc.

Wicherts, J. M. (2016). The importance of measurement invariance in neurocognitive ability testing. *The Clinical Neuropsychologist, 30*, 1006–1016. https://doi.org/10.1080/13854046.2016.1205136.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Jason R. Parkin Ph.D., NCSP is Assistant Clinical Professor at Seattle University. He teaches courses in school psychology and psychoeducational assessment.

Craig L. Frisby Ph.D., Associate Professor of School Psychology at the University of Missouri, Columbia. He has expertise in test session behavior, applications of multidimensional scaling, and the measurement of critical thinking

Ze Wang Ph.D., is Associate Professor in the Department of Educational, School, and Counseling Psychology at the University of Missouri. She has expertise in statistical modeling, measurement invariance, and scale development.