Title

What is text reading fluency and is it a predictor or an outcome of reading comprehension? A longitudinal investigation.

Journal

Developmental Psychology, 57(5)

Authors

Kim, Young-Suk
Quinn, Jamie
Petscher, Yaacov

Publication Date

2021-05-01

DOI

10.1037/dev0001167

Peer reviewed

HHS Public Access
Author manuscript
Dev Psychol. Author manuscript; available in PMC 2022 May 01.

Published in final edited form as:
Dev Psychol. 2021 May ; 57(5): 718–732. doi:10.1037/dev0001167.

What is text reading fluency and is it a predictor or an outcome of reading comprehension? A longitudinal investigation

Young-Suk Grace Kim,
University of California, Irvine

Jamie M. Quinn,
Florida Center for Reading Research

Yaacov Petscher
Florida State University

Abstract

Text reading fluency refers to the ability to read connected texts with accuracy, speed, and expression (prosody), and has garnered substantial attention as an important skill for reading comprehension. However, two fundamental questions remain—the dimensionality of text reading fluency including text reading efficiency (accuracy & speed) *and* reading prosody, and the directionality of the relation between text reading fluency and reading comprehension. These questions were addressed using longitudinal data from Grade 1 (mean age = 6.36 years) to Grade 3 (mean age = 8.34 years). Majority of children were White (approximately 60%) and African American (26%) with 39% to 52% from low SES backgrounds, depending on the grade. Text reading fluency, word reading, listening comprehension, and reading comprehension were measured. Results from confirmatory factor analysis revealed that text reading fluency is a multidimensional construct with a trifactor structure, which has a general factor that captures common ability across text reading efficiency and reading prosody as well as local and specific factors that are unique beyond the general factor. However, the general factor was the most reliable factor, whereas local and specific factors were not reliable. The directionality of the relation between text reading fluency and reading comprehension was addressed by examining two competing structural equation models—text-reading-fluency-as-a-predictor/mediator model and text-reading-fluency-as-an-outcome model—and data supported the former. These results indicate that text reading fluency is a multidimensional construct, and it acts as a predictor, mediating the relations of word reading and listening comprehension to reading comprehension.

Keywords

Reading fluency; oral reading fluency; text reading fluency; reading prosody; reading comprehension; mediation

---

Text reading fluency, known as reading fluency or oral reading fluency, is typically characterized as one's ability to read connected texts with accuracy, speed, and expression

Corresponding Author Young-Suk Grace Kim, Professor, University of California, Irvine, Youngsk7@uci.edu.

(or reading prosody; National Institute of Child Health and Human Development [NICHD], 2000; see Kuhn, Schwanenflugel, & Meisinger, 2010; Kuhn & Stahl, 2003; Wolf & Katzir-Cohen, 2001). Text reading fluency has received substantial attention in research and is regarded as an important, necessary skill for reading comprehension (Fuchs et al., 2001; Kim, 2020; Kuhn et al., 2010; NICHD, 2000; Wolf & Katzir-Cohen, 2001). Not surprisingly, text reading fluency is widely assessed in schools and is explicitly specified as one of the skills to develop in the Common Core State Standards (National Governors Association Center and Council of Chief State School Officers, 2010) and similar state standards in the US. However, a couple of fundamental questions about text reading fluency still remain, that is, the dimensionality of different aspects of text reading fluency (accuracy, speed, and expression) and the directionality of the relation between text reading fluency and reading comprehension.

Before the review of the relevant literature, it is important to clarify terms used in the present study. Instead of the widely used 'reading fluency' or 'oral reading fluency' terms, we use the term 'text reading fluency' to refer to a construct that includes accuracy, speed, and prosody of text reading, and the term 'text reading *efficiency*' to refer to the accuracy and speed of text reading that does not include reading prosody. The term *text* reading fluency indicates that the focal construct or skill is a text-level reading skill, not lexical-level reading skill (word reading fluency or efficiency), and is not limited to an oral mode (although measuring text reading fluency in oral reading is much easier). The distinction between text- versus word-level fluency is important for theoretical precision and for operationalization in empirical work – text reading fluency and word reading fluency are closely related as text reading fluency is an outcome of word reading fluency, but they are dissociable skills because text reading fluency involves post-lexical comprehension processes that word reading does not (see Jenkins et al., 2003; Kim & Wagner, 2015; Wolf & Katzir-Cohen, 2001).

Text Reading Fluency and Reading Comprehension

Text reading fluency has been studied in two lines of research – one that focuses on text reading efficiency (accuracy and speed of text reading) and the other on reading prosody. The role of text reading efficiency in reading comprehension is explained by the automaticity theory (LaBerge & Samuels, 1974) and verbal efficiency theory (Perfetti, 1992): Automatic text reading (operationalized as accurate and rapid text reading) allows cognitive resources to be available for meaning-construction processes, whereas inaccurate and slow reading limits availability of cognitive resources, hampering reading comprehension (Perfetti, 1992). By now a large body of studies has shown strong relations of text reading efficiency to reading comprehension, although the relation is typically stronger for children in elementary school grades than in secondary schools (e.g., Baker, Park, & Baker, 2012; Jenkins et al., 2003; Kim, Petscher, Schatschneider, Foorman, 2010; Kim, Wagner, & Foster, 2011; Klauda & Guthrie, 2008; Sabatini, Wang, & O'Reilly, 2019; Shinn, Good, Knutson, Tilly, & Collins, 1992). For example, when using latent variables, the correlations between text reading efficiency and reading comprehension were .91 in Grade 1, .80 in Grade 2, .83 in Grade 3, and .79 in Grade 4 (Kim & Wagner, 2015). Shinn et al. (1992) reported correlations from .54 to .62 for third and fifth graders. Furthermore, it was shown that text reading efficiency mediates the relations of word reading and listening comprehension to reading comprehension, and the nature of mediation varies as a function of development (or reading development phases; Kim, 2015; Kim, Park, & Wagner, 2014; Kim & Wagner, 2015).

The other aspect of text reading fluency, reading prosody, is suprasegmental rhythmic and melodic rendering of connected text reading, including pitch, stress, and duration of pause (Dowhower, 1991; Schreiber, 1991). Expressive reading or prosodic reading is characterized as reading texts with greater range and variation in pitch (F0), grouping words in meaningful units or pausing at grammatically relevant junctures, and pausing for appropriate durations (e.g., Benjamin et al., 2013; Binder et al., 2013; Dowhower, 1991; Groen, Veenendaal, & Verhoeven, 2019; Schwanenflugel et al., 2004). Reading prosody is expected to facilitate reading comprehension because it "provides a basic cognitive skeleton that allows one to hold an auditory sequence in working memory" and "assists in maintaining an utterance in working memory until a more complete semantic analysis can be carried out" (Kuhn et al., 2010, p. 237). A body of studies has shown the relation of reading prosody to reading comprehension (e.g., Binder et al., 2013; Calet, Gutierrez-Palma, & Defior, 2015; Daane, Campbell, Grigg, Goodman, & Oranje, 2005; Groen et al., 2019; Klauda & Guthrie, 2008; Miller & Schwanenflugel, 2008; Paige, Rasinski, Magpuri-Lavell, & Smith, 2014; Schwanenflugel et al., 2004; Veenendaal, Goen, & Verheoeven, 2014). Klauda and Guthrie (2008), for example, measured reading prosody with a rating scale and reported strong relations with reading comprehension for English-speaking fifth graders (.68 ≤ *r*s ≤ .75; also see Benjamin et al., 2013). Also using a rating scale (i.e., adapted Multidimensional Fluency Scale; Rasinski, 2004), Groen et al. (2019) found a moderate relation (.52) for Dutch-speaking poor comprehenders in Grade 5; similar relations were reported for Spanish-speaking children in Grades 2 and 3 (Calet et al., 2015). Furthermore, reading prosody as measured by spectrographic analysis was related to reading comprehension, but the magnitudes of relations varied depending on prosody indicators. Specifically, pause structure (e.g., inappropriate or ungrammatical pauses) had moderate to strong negative relations with reading comprehension (Binder et al., 2013), whereas pitch-related prosody indicators (e.g., child-adult pitch [F0] match, or sentence-final F0 change) had positive but weak relations with reading comprehension (Binder et al., 2013; Schwanenflugel et al., 2004).

Although prior research on text reading fluency has been highly informative, there are a couple of important gaps in the literature. First, no prior work has investigated its dimensionality – whether text reading efficiency and reading prosody are unidimensional (i.e., text reading efficiency and reading prosody are described by a single underlying skill), multidimensional as two related but dissociable skills, or multidimensional with a bifactor or trifactor structure that is composed of a general factor that captures commonality between text reading efficiency and reading prosody, and local and specific factors of text reading efficiency and reading prosody that are unique beyond the general factor (see the Data Analytic Strategies for details). Previous studies indicate that text reading efficiency and reading prosody are related (e.g., Benjamin et al., 2013; Paige et al., 2014), but may have differential relations depending on the aspects of reading prosody. For example, pause structure indicators of reading prosody such as ungrammatical pauses (intra-sentential or intra-word pauses) are moderately to fairly strongly related with text reading efficiency

(Binder et al., 2013), whereas pitch-related prosody indicators (e.g., sentence-final pitch changes) have little or weak relations with text reading efficiency (Binder et al., 2013; Miller & Schwanenflugel, 2008).

A second gap in the literature is the directionality of the relation between text reading fluency and reading comprehension – whether text reading fluency is a predictor or a product/outcome of reading comprehension. Theoretical accounts state that text reading fluency involves post-lexical *comprehension* processes (see Jenkins et al., 2003; Wolf & Katzir-Cohen, 2001), which help readers anticipate words in connected texts (Jenkins et al., 2003; Perfetti, Goldman, & Hogaboam, 1979). The vast majority of previous studies took this as an explanation for why text reading fluency is strongly related to reading comprehension, and examined the relation of text reading fluency to reading comprehension. However, the directionality deserves more scrutiny: If text reading fluency involves comprehension processes, is text reading fluency necessary for or a predictor of reading comprehension or is it an outcome or by-product of reading comprehension? Theoretical accounts and explanations have been somewhat ambiguous and elusive about this point. Kim (2015), however, posited that comprehension captured in text reading fluency is that of automatic processes related to semantic network, and thus involves low-level comprehension processes needed for comprehending local propositions, not high-level inferential and integration processes that are needed for deep reading comprehension. Results supported the speculation such that text reading efficiency was predicted by vocabulary and grammatical knowledge, but not by higher order cognitive skills such as making inferences about others' viewpoints (i.e., perspective taking). In contrast, reading comprehension was uniquely predicted by perspective taking in addition to vocabulary and grammatical knowledge (Kim, 2015).

Previous studies have found support for both directional relations: Numerous studies have shown the relation of text reading efficiency to reading comprehension (see above) while studies also found support for the relation of reading comprehension to text reading efficiency (Baker, Stoolmiller, Good, & Baker, 2011; Jenkins et al., 2003; Kim, 2015). To our knowledge, only two studies attempted to directly compare alternative models (text reading fluency → reading comprehension versus reading comprehension → text reading fluency) to explicitly examine directionality of the relation between text reading fluency and reading comprehension, and each examined text reading efficiency and reading prosody. Kim (2015), using text reading efficiency, examined two competing models: In one model, text reading efficiency was a predictor/mediator in the relations of word reading and listening comprehension to reading comprehension, and in the other model, text reading efficiency was an outcome of reading comprehension (i.e., reading comprehension was a predictor/mediator in the relations of word reading and listening comprehension to text reading efficiency). The results revealed that both models fit the data very well, and the differences in model fits were ambiguous with no clear indication about a preferred model. Schwanenflugel et al. (2004) investigated similar alternative models between reading prosody and reading comprehension: In one model, reading prosody was a predictor/mediator in the relation of word reading to reading comprehension, whereas in the other model, reading prosody was an outcome of reading comprehension (i.e., reading comprehension was a predictor/mediator in the relation of word reading to reading prosody).

They found that both models fit the data very well, but preferred the former (reading prosody as a predictor) because the latter model (reading prosody as an outcome) had a collinearity between the predictors, reading comprehension and word reading.

Present Study

Our goal in the present study was to fill the two identified gaps in the literature, using longitudinal data. First, we investigated the dimensionality of text reading efficiency and reading prosody. Although the definition of text reading fluency includes text reading efficiency (accuracy and speed) and reading prosody, to our knowledge, no prior work examined the dimensionality of text reading fluency, including both text reading efficiency and reading prosody—whether they capture a single underlying skill, or multiple dimensions, and if the latter, what structure is supported by data.

Second, using the identified dimensions of text reading fluency, we investigated the directionality of the relation between text reading fluency and reading comprehension by comparing two competing models—text reading fluency as a predictor versus text reading fluency as an outcome. The two extant studies that directly compared the competing models (i.e., Kim, 2015; Schwanenflugel et al., 2004) were inconclusive and limited by cross-sectional data; and each examined text reading efficiency and reading prosody, respectively, not text reading fluency that incoporate both. In the present study, we used longitudinal data from children assessed in the fall and spring of each year in Grades 1 to 3. We examined whether text reading fluency measured at a middle time point (e.g., spring of Grade 1) mediates the relations of word reading and listening comprehension at an earlier time point (e.g., fall of Grade 1) to reading comprehension at a later time point (e.g., fall of Grade 2), or alternatively whether reading comprehension measured at a middle time point mediates the relations of word reading and listening comprehension at an earlier time point to text reading fluency at a later time point. As illustrated in Figure 2, these competing models were fitted using two series of data: one series from Time 1 (fall of Grade 1) to Time 3 (fall of Grade 2; see Figures 2a and 2c) and the other series from Time 3 (fall of Grade 2) to Time 5 (fall of Grade 3; see Figures 2b and 2d; see Data Analytic Strategies section for details). The competing models were fitted using two series of data: one series from Time 1 (fall of Grade 1) to Time 3 (fall of Grade 2) and the other series from Time 3 (fall of Grade 2) to Time 5 (fall of Grade 3; see Data Analytic Strategies section for details). Although data were available from six time points, in the analysis, we utilized five time points, Time 1 to Time 5, for comparability of the above noted two time series—fall-to-fall data rather than mixed time points (e.g., fall to fall vs. spring to spring).

Specific research questions that guided the present study were as follows: (a) What is the dimensionality of text reading efficiency and reading prosody variables for children in primary grades?; and (b) what is the directionality of the relation between text reading fluency and reading comprehension? Is text reading fluency a predictor or an outcome of reading comprehension? We hypothesized that a multidimensional model, specifically, a bifactor or a trifactor structure – a general factor that captures common ability across text reading efficiency and reading prosody skill as well as local and specific factors that capture text reading efficiency and reading prosody – would be supported by the data because of evidence of varying relations between text reading efficiency and reading prosody depending on aspects of reading prosody (Binder et al., 2013; Miller & Schwanenflugel, 2008). We also postulated that both the text-reading-fluency-as-a-predictor/mediator model and the text-reading-fluency-as-an-outcome model would fit the data well, but the former would have a superior fit than the latter, if text reading fluency captures low-level comprehension processes, whereas successful reading comprehension requires higher order processes (Kim, 2015).

Method

Participants

The sample comprised of 371 English-speaking children measured in the fall and spring of Grade 1 (fall mean age = 6.36 years [$SD = 0.53$]), Grade 2 (fall mean age = 7.33 years [$SD = 0.52$]), and Grade 3 (fall mean age = 8.34 years [$SD = 0.54$]). Students were from seven schools in two school districts in a large southeastern state in the US. The longitudinal study included measures beyond those used in the present study. For example, results from Grade 1 that focus on emergent literacy skills and reading skills (Kim & Petscher, 2016) and eye movements and reading proficiency (Kim, Petscher, & Vorstius, 2019) were reported earlier. Fifty-two percent of students in Grade 1 ($n = 192$), 46% of students in Grade 2 ($n = 172$), and 39% of students in Grade 3 ($n = 146$) qualified for free or reduced lunch, a proxy variable for low socioeconomic status. The racial/ethnic breakdown was as follows in Grade 1: 59.8% White, 25.9% Black, 5.9% Hispanic, 2.4% Asian/Pacific Islander, and 0.3% American Indian/Alaska Native, and 5.9% identified as two or more races/ethnicities. A little less than half of the sample was female ($n = 180$; 48.3%). Ten students had native languages other than English (Chinese, Hindi, Norwegian, Spanish, Telugu, Turkish), and 15 students reported that English was not the predominant language spoken at home by their parents or guardians. Only three students were identified to have limited English proficiency (LEP) status in Grade 1. Twenty-nine students had a speech-related primary exceptionality in Grade 1, and an additional seven students had one other primary exceptionality (e.g., autism, intellectual disability, learning disability, language impairment). All students were included in the analysis. Human Subjects Review was approved by the Florida State University (HSC No. 2015.16488): Development of oral and silent reading fluency and its relation with reading comprehension in first through third grade students.

Sample sizes were $n = 350$ in Time 1, $n = 366$ in Time 2, $n = 326$ in Time 3, $n = 329$ in Time 4, and $n = 311$ in Time 5. By Time 5, attrition reached 11% in the sample. The number of participants with missing data in each time point were as follows: Time 1: $n = 21$, Time 2: $n = 5$, Time 3: $n = 45$, Time 4: $n = 42$, and Time 5: $n = 60$. Missing data occurred for several reasons, the most likely of which was for technical reasons (e.g., malfunction of digital recorder for recording of oral reading). Pause durations between sentences and F0 changes across sentences were more likely to be missing for the second and third passages in Time 1 and Time 2. Additionally, some students experienced glottal fry (i.e., creaky voice) during the recording of their oral reading, so their recordings were not coded. There were no statistically significant differences in students who had missing data versus those students who did not on any demographic variables according to chi-square difference tests (*ps* > .10). One exception was in Time 1, where students who qualified for free or reduced lunch were more likely to have missing data than students who did not qualify for free or reduced lunch (χ^2 [1] = 5.03, p = .025, Cramer's V = .12). More information on these chi-square difference tests are available in the online supplemental materials. We conducted Little's test of missing data, which tests whether data can be considered Missing Completely at Random (MCAR), on the raw data within each time point. The tests failed to establish MCAR within the data for Time 2 to Time 4 (all *ps* < .001). Data were considered MCAR in Time 5 (χ^2 (4097) = 4049.04, p = .700). We considered the data to be "Missing at Random" (MAR) in Time 2 through Time 4 due to issues with spectrographic measurement at these times (see below).

Measures

Students' text reading efficiency, reading prosody, word reading, listening comprehension, and reading comprehension skills were measured at each time point. However, aligned with the research questions and associated data analytic strategies, data used in the present study were as follows: Text reading efficiency, reading prosody, and reading comprehension at Times 2, 3, 4, and 5, and word reading and listening comprehension at Times 1 and 3. Unless otherwise noted, all items were dichotomously scored, and reliability estimates are from the present sample.

Text reading efficiency.—Three grade-level passages were presented to children at Times 2, 3, 4, and 5, and students were asked to read each of the three passages aloud as best as they could. The passages were part of the Florida Center for Reading Research Reading Assessment (Foorman, Petscher, & Schatschneider, 2015), which were normed in the State of Florida. The passages were equated, and were composed of narrative and expository texts (4 out of 9 passages in the present study were narratives). Grade 1 passages had 155 to 198 words (400 to 700 Lexiles), Grade 2 passages had 187 to 200 words (600 to 780 Lexiles), and Grade 3 passages had 200 to 307 words (400 to 790 Lexiles). One passage at each time point served as the linking passage between time points (i.e., one passage at Time 2 was also given at Time 3, etc.) and thus a total of 9 different passages were used across the four time points. Students' score for each passage was the number of words read accurately in 1 minute. Parallel form reliability exceeded .89 across the three passages within each time point (Time 2: $.91 \leq r \leq .92$; Time 3: $.90 \leq r \leq .92$; Time 4: $.91 \leq r \leq .93$; Time 5: $.90 \leq r \leq .92$). Test-retest reliability ranged from .87 to .88 for linking passages.

Reading prosody.—Students' reading prosody was measured from the text read aloud for text reading efficiency. Students' oral reading was digitally recorded, and digital recordings (*.wav files) were used in spectrographic analysis and a ratings scale. Spectrographic evaluation included two measures of pitch – vocalic nucleus and sentence-final change in F0 – and three measures of pause structure – inter-sentential pause duration, ungrammatical pause frequency, and total pause frequency. Details of each are found in Table S1 in the online supplemental materials.

Due to the resource-intensive nature of the coding and the large amount of data (e.g., for 300 students at each time point, there were approximately 2700 sentences [3 sentences per passage*3 passages*300 students] to code for multiple aspects using spectrogram), the first three sentences from the digital recording of each passage were used for spectrographic analysis (see Miller & Schwanenflugel, 2008; Schwanenflugel et al., 2004, for a similar approach). One graduate student and three undergraduate students in the speech and language pathology program underwent rigorous training and coded data using Praat software (Version 5.4; Boersma & Weenink, 2015). Similarity reliability coefficients (Shrout & Fleiss, 1979), which indicated the proximity of the coder's score for each variable to that of the primary coder, ranged from .90 to .99, using 78 cases. The training was composed of several sessions where research assistants were introduced to focal prosody features and practiced coding them using the Praat software. This was followed by additional practice sessions until research assistants indicated their readiness to try a reliability set. If reliability was not met in the initial trial (e.g., a minimum of .80), differences were discussed, and subsequent reliability sets were conducted until the minimum reliability was met in each prosody feature.

Reading prosody was also evaluated using a widely used ratings scale, the Multi-Dimensional Fluency Scale (Rasinski, 2004). The Multi-Dimensional Fluency Scale scoring guide includes four aspects: (a) expression and volume, (b) phrasing, (c) smoothness, and (d) pacing. Each aspect was rated on a scale of 1 to 4 (1 for not fluent reading to 4 for fluent reading). Three individuals (one doctoral student in Education, one individual with a master's degree in Education, and one individual with a bachelor's degree in speech and language pathology) were rigorously trained with exact percent agreement ranging from .80 to .90, using 72 cases. The training followed a very similar approach described above for the spectrographic analysis.

Word reading.—Word reading skill was assessed with three tasks: the Letter-Word Identification (LWID) task of the Woodcock-Johnson III Tests of Achievement (WJ-III; Woodcock, McGrew, & Mather, 2001); the Word Reading subtask of the Wechsler Individual Achievement Test, 3rd edition (WIATWR; Wechsler, 2009); and the Sight Word Efficiency (SWE) subtask of the Test of Word Reading Efficiency, 2nd edition (Torgesen, Wagner, Rashotte, 2012). LWID and WIATWR were untimed tasks, where the student was asked to read aloud from a list of words of increasing difficulty with accuracy. SWE was a timed task where the student was asked to read a list of words of increasing difficulty as accurately as possible in 45 seconds. Cronbach's alpha estimates were .91 to .92 in LWID and .95 in WIATWR at each time point. Test-retest reliability for SWE was reported to be .93 (Torgesen et al., 2012).

Listening comprehension.—Listening comprehension was measured by two normed tasks, the Oral Comprehension subtest of WJ-III (WJOC henceforth; Woodcock et al., 2001) and the Listening Comprehension subscale of the Oral and Written Language Scales, 2nd Edition (OWLS henceforth; Carrow-Woolfolk, 2011). In WJOC, the student was presented a spoken sentence or short paragraph and was asked to supply a missing word. In OWLS, the student was presented with sentences read aloud and was asked to point to one of four pictures that best depicted the spoken stimulus. Cronbach's alpha estimates ranged from .74 to .79 in WJOC and .91 to .93 in OWLS.

Reading comprehension.—Reading comprehension was assessed with two normed tasks, the Reading Comprehension subtest of the WIAT (WIATRC henceforth; Wechsler, 2009) and the Passage Comprehension subtest of the WJ-III (WJPC henceforth; Woodcock et al., 2001). In WIATRC, students were asked to read various types of texts (e.g., fiction, informational, how-to) and respond to literal and inferential questions about the passage. In WJPC, a cloze task, the student was asked to supply a missing word to sentences and paragraphs of increasing complexity. Cronbach's alpha estimates were .86, .87, .85, and .86 in WIATRC and .81, .86, .75, and .84 in WJPC at Times 2, 3, 4, and 5, respectively.

Procedures

Children were individually assessed by rigorously trained research assistants in a quiet space in participating schools. The assessments were administered in several sessions (each session taking approximately 30 to 40 minutes).

Data Analytic Strategies

Confirmatory factor analysis (CFA) and structural equation modeling (SEM) were conducted using Mplus version 8.4 (Muthén & Muthén, 1998–2019). All models were fit using maximum likelihood estimation with robust standard errors (MLR) to account for the censored nature of the pause variables, where both durations and number of occurrences were censored from below (i.e., had values close to zero).

Research question 1: Dimensionality of text reading efficiency and reading prosody variables.—The dimensionality of the prosody indicators and the text reading efficiency indicators was determined within each time point using CFA. For text reading efficiency (i.e., words read correctly per minute within a passage), three separate indicators for the three text reading efficiency passages were used in the dimensionality models. For reading prosody, pause frequencies and ungrammatical pause frequencies were averaged due to extremely high correlations ($.98 \leq r$s $\leq .99$). Therefore, for reading prosody, there were four indicators for spectrograph data—intonation contour (i.e., vocalic nuclei), F0 change, pause frequencies, and pause durations—and four indicators of the ratings scale—expression and volume, phrasing, pacing, and smoothness—with each indicator averaged across passages per time point.

Based on theory and evidence from prior work, we fit six plausible alternative CFA models to examine the dimensionality of the reading prosody and text reading efficiency data (see Figure 1). First, we fit a unidimensional model (Figure 1a), whereby one factor was indicated by the eleven variables representing reading prosody and text reading efficiency (i.e., intonation contour, F0 change, pause frequencies, pause durations, expression and volume, phrasing, pacing, smoothness, and the words correct per minute for the three text reading efficiency passages). The second model, Figure 1b, specified two factors, one representing reading prosody and a second factor representing text reading efficiency using the words correct per minute of the three passages as indicators. The third model, Figure 1c, retained the two factors from the second model, Figure 1b, but instead fit a bifactor structure, such that a general factor represented the common variance among the eleven indicators along with a reading prosody specific factor and text reading efficiency specific factor. In a bifactor model (Gibbons & Hedeker, 1992), the general factor (in this case, Text Reading Fluency) captures common variance across all the manifest variables or indicators, and thus, it theoretically captures the most reliable portion of the variance for each of the indicators. The specific factors, orthogonal to the general factor and to each other, help to explain item response variance (item residual variance) that is not captured by the general factor (e.g., method variance).

In the next three models, Figures 1d, 1e, and 1f, the ratings and pause-structure reading prosody indicators had a bifactor structure, whereas pitch-related prosody indicators (i.e., intonation contour and F0 change) formed their own factor as did text reading efficiency indicators (i.e., passage 1, passage 2, and passage 3). That is, the ratings and pause-structure indicators had a general factor, called Prosody: Ratings and Pause, which captures common variance among the four ratings variables (i.e., expression and volume, phrasing, pacing, and smoothness) and the two pause variables (i.e., pause durations and pause frequencies), along with a Ratings specific factor and a Pause specific factor. Prior evidence suggested that pitch and pause structure may be related but tap into different aspects of reading prosody (see Benjamin et al., 2013) as pitch variables have weak relations with pause variables (Binder et al., 2013; Schwanenflugel et al., 2004). Furthermore, our work without text reading efficiency variables supported a bifactor structure for the ratings variables and pause-structure prosody variables (Kim, Quinn, & Petscher, 2020).

In the fourth model, Figure 1d, correlations were allowed between Text Reading Efficiency, Prosody: Ratings and Pause, and Prosody: Pitch factors. The specific Ratings factor and specific Pause factor were not allowed to correlate with other factors in the model. The fifth model (Figure 1e) took the structure of Figure 1d and added a second order factor, where the Prosody: Ratings and Pause, Prosody: Pitch, and Text Reading Efficiency latent variables loaded onto a higher order factor named Text Reading Fluency. The final model (Figure 1f) was a trifactor model. Trifactor models are conceptually similar to bifactor models. Bifactor models traditionally specify one global, latent factor causing all indicators and two or more specific, latent factors causing an orthogonal set of indicators; thus, each indicator loads on two latent factors. Trifactor models extend this logic to allow combinations of measured variables to load onto *three* types of latent factors: a general factor, a specific factor, and a local factor. General and specific factors retain their same interpretation as in the bifactor model. A local factor is a theoretical aggregate of content-specific areas. As shown Figure 1f, there were three local factors (Ratings and Pause, Prosody: Pitch, and Text Reading Efficiency) and two specific factors (Ratings, Pause) in addition to the general factor. Text Reading Fluency is a general factor that included all measured variables as indicators (i.e., expression and volume, phrasing, pacing, smoothness, pause durations, pause frequencies, intonation contour, F0 change, and the three text reading efficiency variables). One local factor, Prosody: Ratings and Pause, captured variances in expression and volume, phrasing, pacing, smoothness, pause durations, and pause frequencies over and above the general factor. The specific Ratings factor captured variance in the ratings variables over and above the general factor and the local factor, Prosody: Ratings and Pause; the specific Pause factor captured variance in pause durations and pause frequencies over and above the general factor and the local factor, Prosody: Ratings and Pause. The Prosody: Pitch local factor captured variance in intonation contour and F0 change over and above the general factor, and the Text Reading Efficiency local factor captured variance in the three text reading efficiency variables over and above the general factor.

Research question 2: Directionality of the relation between text reading fluency and reading comprehension.—After establishing the dimensionality of the text reading efficiency and reading prosody variables, we examined the relations among word reading, listening comprehension, text reading fluency, and reading comprehension using SEM. First, the following latent variables were created using confirmatory factor analysis: word reading, indicated by the three normed word reading measures (i.e., LWID, WIATWR, and SWE); listening comprehension, indicated by the two normed language measures (i.e., WJOC and OWLS); and reading comprehension, indicated by the two normed comprehension measures (i.e., WIATRC and WJPC). We then tested the two competing mediation models shown in Figure 2 based on theoretical accounts (Kim, 2015; Kuhn et al., 2010; Wolf & Katzir-Cohen, 2001) and empirical evidence (Kim, 2015; Kim & Wagner, 2015; Schwanenflugel et al., 2004).

Model 1a and Model 1b: Text reading fluency as the mediator.—In this model, we used text reading fluency as a mediator in the relation of word reading and listening comprehension to reading comprehension (see Figure 2a [Model 1a] and Figure 2b [Model 1b]). Model 1a used Time 1 word reading and listening comprehension (fall of Grade 1), Time 2 text reading fluency (spring of Grade 1), and Time 3 reading comprehension (fall of Grade 2), while Model 1b used Time 3 word reading and listening comprehension (fall of Grade 2), Time 4 text reading fluency (spring of Grade 2), and Time 5 reading comprehension (fall of Grade 3). The models accounted for baseline levels of text reading fluency and reading comprehension (see Figure S2a in the online supplementary materials).

Model 2a and Model 2b: Reading comprehension as the mediator.—In these models, the reading comprehension latent variable was a mediator in the relations of word reading and listening comprehension to text reading fluency (see Figure 2c [Model 2a] and Figure 2d [Model 2b]). Model 2a used Time 1 word reading and listening comprehension (fall of Grade 1), Time 2 reading comprehension (spring of Grade 1), and Time 3 text reading fluency (fall of Grade 2), while Model 2b used Time 3 word reading and listening comprehension (fall of Grade 2), Time 4 reading comprehension (spring of Grade 2), and Time 5 text reading fluency (fall of Grade 3). The models also accounted for baseline levels of text reading fluency and reading comprehension (see Figure S2b in the online supplementary materials).

Non-nested model comparison for Model 1a/2a and Model 1b/2b.—Model 1a and Model 1b, where text reading fluency was the mediator, cannot be directly compared to Model 2a and Model 2b, where reading comprehension was the mediator, as these models had different outcome variables of interest (e.g., Model 1a had Time 3 reading comprehension as the outcome, whereas Model 2a had Time 3 text reading fluency as the outcome). Therefore, we conducted fully saturated analyses where all variables were included in the model (i.e., reading comprehension and text reading fluency at both time points were included). These models can then be compared using the Bayesian Information Criterion (BIC). In the two test models (Test 1 and Test 2; see Figure S1 and associated explanation in the online supplemental materials for further information on these test models), the same sets of variables are used: Time 1 listening comprehension and word reading, Time 2 text reading fluency and reading comprehension, and Time 3 text reading fluency and reading comprehension. The difference between these two test models is which construct is included as a mediator. In Test 1, text reading fluency was included as the mediator. Therefore, all pathways involving Time 2 reading comprehension and Time 3 text reading fluency were fixed to zero. In Test 2, reading comprehension was included as the mediator, and so all pathways involving Time 2 text reading fluency and Time 3 reading comprehension were fixed to zero. Test 3 and Test 4 were conducted in the same way, using variables from Time 3, Time 4, and Time 5. The resulting BIC values were then compared to determine the better fitting model between Test 1 and Test 2 and between Test 3 and Test 4.

Model fit.—Multiple criteria were used to determine model fit. The confirmatory fit index (CFI) and the Tucker-Lewis Index (TLI) were used, whereby values above .90 are considered adequate, and values above .95 are considered excellent (Hu & Bentler, 1999). We also used the root mean squared error of approximation (RMSEA) and its associated confidence interval, where values under .08 are preferred (Kline, 2016). For determining the best fitting nested models, the difference in the Satorra-Bentler chi-square tests of model fits were used, whereby the preferred model was one in which the difference in Satorra-Bentler chi-square estimates was significantly better, or significantly closer to zero (Muthén & Muthén, 1998–2018). When an assumption of nested models was not met (e.g., comparing Figure 1e to Figure 1f), we looked at the differences in the sample-size adjusted Bayesian Information Criterion (*n*BIC) between compared models. Models with *n*BIC values closer to negative infinity were preferred, where a difference of 5 was considered *strong* evidence for a better fitting model and a difference of 10 was considered *very strong* evidence for a better fitting model (Raftery, 1995).

Results

Descriptive Statistics

Descriptive statistics for text reading efficiency, reading prosody, and reading comprehension from Time 2 to Time 5 are presented in Table 1. Text reading efficiency scores increased over time. Intonation contour means decreased from Time 2 (258.19) to Time 5 (247.2). Overall, pause durations and pause frequencies decreased over time. As for reading prosody measured by the ratings scale, average scores on the four ratings scale categories increased over time. As mentioned above, reading prosody data from spectrographic analysis and the ratings scale were averaged across passage. Descriptive statistics for the spectrographic and ratings scale variables according to each passage are presented in the online supplemental materials (Table S2). The average standard scores of WIATRC and WJPC indicated that the sample was in the average range at all time points.

Descriptive statistics for the word reading and listening comprehension measures at Times 1 and 3 are presented in Table 2.

Correlation matrices for the concurrently measured prosody and text reading efficiency variables are presented in Table 3 (see the top panel for Time 2 and Time 3, and the bottom panel for Time 4 and Time 5). There were strong correlations among text reading efficiency variables at all time points (rs ≥ .90). Text reading efficiency variables were negatively related to F0 change (−.26 ≤ r ≤ −.17), pause durations (−.59 ≤ r ≤ −.51), and pause frequencies (−.87 ≤ r ≤ −.77). Ratings scale indicators were strongly and positively correlated with text reading efficiency scores (.70 ≤ r ≤ .86). Intonation contour was negatively correlated with F0 change (−.46 ≤ r ≤ .35), and was weakly correlated with other variables (−.10 ≤ r ≤ .19). The ratings scale variables were negatively correlated with F0 change (−.40 ≤ r ≤ −.16), pause durations (−.54 ≤ r ≤ −.44), and pause frequencies (−.80 ≤ r ≤ −.70). There were strong, positive correlations between ratings scale indicators (.78 ≤ r ≤ .89). Longitudinal correlation matrices used for the structural equation models are presented in Tables S3 and S4 in the online supplemental materials. Text reading efficiency and reading prosody variables had little (r = .03) to strong (r = .74) relations with reading comprehension.

Research Question 1: Dimensionality of Text Reading Fluency

Six alternative models shown in Figure 1 were fit to the data to determine the dimensionality of the prosody and text reading efficiency indicators. Model fit statistics for the data per time point are presented in Table S5 in the online supplemental materials. The unidimensional model (Figure 1a) fit poorly across all time points, with high chi-square values and poor RMSEAs. Figure 1c fit better than Figure 1b across all time points (Δ*n*BIC range = 10.58–103.52). Figure 1d fit better than Figure 1c at each time point (Δ*n*BIC range = 27.38–155.22). Figure 1d was preferred over Figure 1e in Time 2 (Δ*n*BIC = 47.95), but Figure 1d and Figure 1e were indiscernible in Time 3 (Δ*n*BIC = 0.53), Time 4 (Δ*n*BIC = 2.27), and Time 5 (Δ*n*BIC = 3.90). However, Figure 1f was preferred over both Figure 1d and Figure 1e across all times, with *n*BIC differences of a minimum of 5.28 (considered a practically significant difference in fit) to a maximum of 67.28.

Model parameter estimates, standard errors, and *p*-values for Figure 1f, separated by time point, are presented in Table S6 in the online supplemental materials. Across all four time points, the explained common variance (ECV; Reise, 2012) for the text reading fluency general factor was between .74 and .76, indicating that approximately 75% of the *common* variance was attributable to the general factor (or 25% of the *common* variance was due to the other factors). Coefficient omega (ω; McDonald, 1999) was .91, .92, .92, and .90 for Time 2, Time 3, Time 4, and Time 5, respectively. This suggests that the general factor was very reliable, or that a multidimensional composite score using general factor scores would provide a reliable estimate of text reading fluency. Coefficient omega hierarchical (ω_H; Zinbarg, Revelle, Yovel, & Li, 2005) was .85, .84, .83, and .83, respectively. This indicated that approximately 83–85% of the variance in the multidimensional composite score can be attributed to the general factor, and approximately 90–93% of the *reliable* variance in the multidimensional composite scores is attributable to the general factor (calculated by dividing omega hierarchical ω_H by omega ω; Reise, 2012).

The ECV value ranges for the specific factors were .04-.06 for Ratings and .05-.07 for Pause; and the ranges for the for the local factors were .02-.04 for Prosody: Ratings and Pause, .10-.12 for Prosody: Pitch, and .01-.03 for Text Reading Efficiency. Omega subscale, ω_S, can be calculated and interpreted much in the same way as omega hierarchical ω_H (i.e., a measure of specific factor or local factor reliability). Across all subscales, no ω_S value reached above .09, indicating that these local and specific factors are unreliable and uninterpretable without the context of the general factor. Therefore, only the general factor will be interpreted and explained below for each time point.

Across all four time points, the ratings scale variables consistently and strongly loaded onto the general factor (.77 ≤ λs ≤ .90). Pause frequencies and pause durations negatively loaded onto the general factor across all four time points, with pause frequencies strongly loading (−.91 ≤ λs ≤ −.82) and pause durations moderately but significantly loading (−.58 ≤ λs ≤ −.57). F0 change significantly and negatively loaded onto the general factor, but the magnitude was small (−.27 ≤ λs ≤ −.18). Text reading efficiency scores loaded strongly and positively at all four time points (.91 ≤ λs ≤ .97). The only indicator that did not load significantly onto the general factor at any time point was intonation contour (.06 ≤ *p*s ≤ .84).

Research Question 2: Directionality of the Relation between Text Reading Fluency and Reading Comprehension

We used the identified dimensions of text reading fluency—the Figure 1f model—in two sets of mediation analyses shown in Figure 2. In these models, the general factor for text reading fluency, the most reliable portion of the dimensionality modeling from RQ1, was used as the mediator or outcome. The measurement model of text reading fluency, where the local factors of prosody: ratings and pause, prosody: pitch, and text reading efficiency, and specific factors of ratings and pause, were included in the mediation models reported below, but was not shown in the accompanying figures (Figures 2 and 3) for simplicity.

Model fit statistics for Model 1a, Model 1b, Model 2a, and Model 2b are presented in the top panel of Table S7, and the model comparison tests, Test 1 through Test 4, are presented at the bottom of Table S7 in online supplemental materials (see Figure S1 in the online supplemental materials for more details about the model comparison tests). The BIC for Test 1 (53183.8) was *lower* than the BIC for Test 2 (53316.3). This difference (ΔBIC = 132.5) preferred Test 1, where text reading fluency was set as the mediating variable between word reading and listening comprehension and the outcome variable of reading comprehension (Model 1a shown in Figure 2a). Similarly, Test 3 (49423.8) had a *lower* BIC value than Test 4 (49680.6). This difference (ΔBIC = 256.8) also preferred the model where text reading fluency was set as the mediator (i.e., Test 3; Model 1b shown in Figure 2b). Standardized coefficients from Model 1a and Model 1b, with text reading fluency as the mediator, are presented in Figure 3.

In Model 1a (Figure 3a), word reading moderately and positively predicted text reading fluency (γ = .42, p < .001). There was a weak but statistically significant positive prediction from listening comprehension to text reading fluency above and beyond the effects of word reading (γ = .09, p = .03). Text reading fluency positively and strongly predicted reading comprehension (β = .29, p = .002) above and beyond the relations of listening comprehension (γ = 27, p < .001) and word reading (γ = −.33, p < .001) to reading comprehension. There was evidence of suppression in this model, whereby a relation between two latent variables was negative but the model correlation was positive. Specifically, the relation between word reading and reading comprehension was negative (γ = −.33), but the bivariate correlation between these two latent variables was positive (r = .77, p < .001). Overall, approximately 97% of the variance in reading comprehension and 82% of the variance in text reading fluency were explained.

The results of Model 1b using the later time points of data (i.e., Time 3, Time 4, and Time 5) are shown in Figure 3b. Word reading positively and moderately predicted text reading fluency (γ = .38, p < .001), but listening comprehension (γ = .02, p = .48) did not predict text reading fluency above and beyond word reading and baseline text reading fluency and reading comprehension. Text reading fluency was not related to reading comprehension (β = .11, p = .30) after accounting for word reading (γ = .09, p = .49) and listening comprehension (γ = .23, p = .01) and baseline text reading fluency and reading comprehension. Overall, approximately 100% of the variance in reading comprehension and 91% of the variance in text reading fluency were explained. Full models with baseline controls are found in Figure S2 of the online supplemental materials.

Discussion

In this study, we examined the dimensionality of text reading fluency by including text reading efficiency and reading prosody indicators, and also investigated the directionality of the relation between text reading fluency and reading comprehension, using longitudinal data from Grades 1 to 3.

Dimensionality of text reading fluency

The patterns of bivariate relations between text reading efficiency and various aspects of reading prosody ranged from none (.00) to very strong (.86; see Table 3). Specifically, pitch-related prosody variables (F0 change and intonation contour) had little to weak relations, whereas pause-structure related variables (pause frequency and pause duration) and ratings scale variables had moderate to strong relations with text reading efficiency. Furthermore, pitch-related prosody variables were not strongly related to the other aspects of prosody, pause structure and ratings scale. The strong relations of pause structure variables and weaker relations of pitch-related variables to text reading efficiency were expected, in line with previous findings (Benjamin et al., 2013; Binder et al., 2013; Miller & Schwanenflugel, 2008; Paige et al., 2014). However, the findings of moderate to strong relations of the ratings scale with text reading efficiency is new, which indicates that reading prosody as measured by the ratings scale is likely influenced by decoding skill (bivariate correlations in Tables S3 and S4 of online supplemental materials support this) at least during the beginning phase of reading development examined in the present study. The large constraining role of decoding/word reading skill in reading development has been recognized widely in theoretical models and studies (e.g., Kim, 2020; Hoover & Gough, 1990), and the present study suggests its influence on prosodic rendering of reading texts as well.

Importantly, we found that text reading fluency is best described as a multidimensional structure – specifically a trifactor structure that is composed of a general factor that captures common ability across text reading efficiency and various aspects of reading prosody (e.g., pitch, pause structure), and local factors and specific factors such as text reading efficiency and aspects of reading prosody (ratings, pause structure, pitch) that capture residual variance after accounting for the general factor. Moreover, it was the general text reading fluency factor, not the local or specific factors, that was most reliable and predicted reading comprehension. This was the case across the time points from Grade 1 to Grade 3. To our knowledge, this is the first study that examined the dimensionality of text reading fluency, considering efficiency and prosody of connected text reading. These findings suggest that the widely agreed upon aspects of text reading fluency—accuracy, speed, and prosody of connected text reading—capture a common ability as well as specific abilities. As mentioned above, despite the consensus on these multiple aspects of text reading fluency as a construct (see Fuchs et al., 2001; Kuhn et al., 2010; Wolf & Katzir-Cohen, 2001), the vast majority of studies have operationalized text reading fluency as text reading efficiency, capturing the accuracy and speed aspects, but not reading prosody. Even in studies that included reading prosody, the dimensionality of text reading fluency including text reading efficiency and reading prosody was not addressed.

Nature of the relation between text reading fluency and reading comprehension

Theoretical accounts such as the automaticity theory (LaBerge & Samuels, 1974) explain why accurate and rapid reading is important to supporting reading comprehension. However, the distinctive and defining feature of ‘text’ reading fluency from word reading efficiency (accurate and rapid reading of individual words out of context) is that text reading fluency involves and captures post-lexical comprehension processes (Kim, 2015; Nathan & Stanovich, 1991; Wolf & Katzir-Cohen, 2001). Then, is text reading fluency a predictor or an outcome/by-product of reading comprehension? The vast majority of previous studies assumed the former and examined the relations as such, but only a few studies explicitly examined the directionality of the relation.

In the present study, we examined the directionality question using longitudinal data, and found that text reading fluency, the general factor that captured common ability across text reading efficiency and reading prosody indicators, is not an outcome of reading comprehension, but is a predictor, mediating the relations of word reading and listening comprehension to reading comprehension. This was the case across the two longitudinal series of data (one from the beginning of Grade 1 to the beginning of Grade 2; the other from the beginning of Grade 2 to the beginning of Grade 3). These findings disambiguate previous findings using cross-sectional data (Kim, 2015; Schwanenflugel et al., 2004) and provide further support for the role of text reading fluency in reading comprehension. This finding, rather than the text-reading-fluency-as-an-outcome model, is likely due to the fact that although text reading fluency does capture comprehension processes to some extent, it does not capture high-level or deep comprehension such as global integration processes (Kim, 2015). In other words, text reading fluency does capture reading subcomponent and associated processes, including low-level comprehension processes (e.g., fast-acting automatic semantic network processes), and therefore, is a predictor of reading comprehension – or to be precise it mediates one's word reading skill and comprehension (as captured by listening comprehension). However, successful reading comprehension requires additional high-level inferential and integration processes (e.g., making bridging and elaborative inferences), and therefore, mediation of comprehension processes by text reading fluency is partial (also see Kim, 2015; Kim & Wagner, 2015). Together with a previous study (Kim, 2015), the present findings expand and shed light on the construct of text reading fluency and its role in reading comprehension – previous theoretical conceptualizations did include comprehension processes as part of text reading fluency (see for example, Fuchs et al., 2001; Kuhn et al., 2010; Wolf & Katzir-Cohen, 2001), but were vague about specific differential processes involved in text reading fluency versus reading comprehension.

Although not the primary focus of the present investigation, it is noteworthy that the relations among word reading, listening comprehension, text reading fluency, and reading comprehension across the two sets of time points in this study are somewhat divergent from previous findings. Specifically, in previous findings, text reading fluency (operationalized as text reading efficiency) mediated the relation of word reading to reading comprehension at a developmentally later time point (Kim, 2015; Kim & Wagner, 2015). In the present study, the opposite pattern was found such that text reading fluency mediated the relation of word reading and reading comprehension at an earlier time point, whereas this mediation did not occur at a later time point (see Figure 3). Reasons for the discrepancy between the present findings and those of previous studies is unclear. One potential reason is that the construct of text reading fluency in the present study included reading prosody, whereas previous studies operationalized text reading fluency as text reading efficiency. Another explanation is a difference in data analysis strategy. In the present study, the structural relations were examined longitudinally (i.e., word reading and listening comprehension at an earlier time point predicting text reading fluency at a later point, which in turn predicted reading comprehension at a later time point). In contrast, in previous studies, the structural relations were examined cross-sectionally. Future replications are necessary to further illuminate the nature of developmental relations.

Limitations, Future Studies, and Implications

The findings of the present study are from English-speaking children in primary grades, and therefore, the generalizability of these findings should be further investigated in future studies with children at more advanced developmental phases and with children learning to read in various orthographies. For example, the percentage of students with learning disabilities was small in the present study, and therefore, generalizability may be limited for this population. Another relevant point is that as shown in the loadings in Table S6 in the online supplemental materials, the large loadings for the text reading fluency latent variable were mostly from text reading efficiency variables, pause structure variables, and ratings variables. In contrast, pitch-related variables (F0 change and intonation contour) had weak to little loadings. Therefore, despite the fact that the text reading fluency latent variable did include prosody variables, it appears that the latent variable largely captured word reading and pause-related reading skills in connected texts during the examined developmental points from Grade 1 to Grade 3. Again, these findings are likely due to large constraining roles that word reading/decoding skills play in reading development, including reading prosody, such that the ratings scale and pause structures of reading prosody have large shared variance with text reading efficiency in the beginning phase of reading development. At a more advanced phase of reading development, however, the constraint of word reading is reduced, and consequently the contributions or loadings of pitch-related prosody variables are expected to increase. Then, the construct of text reading fluency at a later developmental phase is likely to be different than that at an earlier developmental phase (which was captured in this study). In other words, the construct of text reading fluency and what is captured in reading prosody may change as a function of reading development. Therefore, it is important to replicate the present study with children at a later reading development phase.

Furthermore, we did not perform a power analysis for the present paper; however, given that the sample and data were part of a larger funded grant project where Monte Carlo simulations were performed for power for detecting single versus multidimensional structures, our questions can be considered adjacent to the questions powered by the initial Monte Carlo study and are sufficiently powered for the present study. Another fruitful direction of future studies is an examination in languages with transparent orthographies, particularly using longitudinal data. Word reading develops at a faster rate in language with transparent orthographies, provided adequate instruction (Seymour, Aro, & Erskine, 2003). As such, developmentally changing relations among text reading efficiency, word reading, listening comprehension, and reading comprehension was shown to occur at a faster rate in transparent orthographies (e.g., Carretti et al., 2020; Kim, 2015; Kim, Park, & Wagner, 2014). Following the same logic, the developmentally changing nature of the construct of text reading fluency that includes reading prosody may be observed at a faster rate in transparent orthographies than in opaque orthographies (e.g., English).

Finally, in the present study, students' reading prosody was measured using the first three sentences of each passage, given the resource-intensive nature and following prior work (e.g., Miller & Schwanenflugel, 2008; Schwanenflugel et al., 2004). Future work can replicate the present study using students' read-aloud of the entire passages to examine the representativeness of students' reading the first 3 sentences compared to entire passage readings.

The finding of the mediating role of text reading fluency suggests that effective instruction on text reading fluency is important to improving reading comprehension. Studies have shown that explicit and systematic instruction on text reading efficiency does causally improve reading comprehension (Vadasy & Sanders, 2008; also see Kuhn & Stahl, 2003, and NICHD, 2000). The present study also indicates the importance of decoding on reading prosody. Then, it is not clear whether instructional attention on reading prosody itself has added values beyond instruction on decoding (and comprehension). To our knowledge, few studies on text reading fluency instruction systematically combined text reading efficiency and reading prosody and examined their impacts on reading comprehension. Future studies are warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported by Grants R305A120147 and R305A180055 by the Institute of Education Sciences, US Department of Education, and by the National Institute of Child Health and Human Development (P50HD052120). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. The authors wish to thank participating schools, teachers, and children.

References

Arcand M-S, Dion E, Lemire-Theberge L, Guay M-R, Barrette A, … Fuchs D (2014). Segmenting texts into meaningful word groups: Beginning readers' prosody and comprehension. Scientific Studies of Reading, 18(3). 208–223. DOI: 10.1080/10888438.2013.864658

Baker DL, Park Y, & Baker SK (2012). The reading performance of English learners in grades 1–3: The role of initial status and growth on reading fluency in Spanish and English. Reading and Writing, 25, 251–281. doi:10.1007/s11145-010-9261-z

Baker DL, Stoolmiller M, Good RH, & Baker SK (2011). Effect of reading comprehension on passage fluency in Spanish and English for second-grade English learners. School Psychology Review, 40, 331–351.

Benjamin RG, & Schwanenflugel PJ (2010). Text complexity and oral reading prosody in young readers. Reading Research Quarterly, 4, 388–404. 10.1598/RRQ.45.4.2

Benjamin RG, Schwanenflugel PJ, Meisinger EB, Groff C, Kuhn MR, & Steiner L (2013). A spectrographically grounded scale for evaluating reading expressiveness. Reading Research Quarterly, 48, 105–133. doi:10.1002/rrq.43

Binder KS, Tighe E, Jiang Y, Kaftanski K, Qi C, & Ardoin SP (2013). Reading expressively and understanding thoroughly: An examination of prosody in adults with low literacy skills. Reading and Writing, 26, 665–680. DOI:10.1007/s11145-012-9382-7 [PubMed: 23687406]

Boersma P, & Weenink D (2015). Praat, doing phonetics by computer. Version 5.4. 20

Calet N, Gutierrez-Palma N, & Defior S (2015). A cross-sectional study of fluency and reading comprehension in Spanish primary grade children. Journal of Research in Reading, 38, 272–285. DOI:10.1111/1467-9817.12019

Carretti B, Toffalini E, Saponaro C, Viola F, & Cornoldi C (2020). Text reading speed in a language with a shallow orthography benefits less from comprehension as reading ability matures. British Journal of Educational Psychology, DOI:10.1111/bjep.12307

Carrow-Woolfolk E (2011). Oral and written language scales (2nd ed.). Torrance, CA: Western Psychological Services.

Daane MC, Campbell JR, Grigg WS, Goodman MJ, & Oranje A (2005). The Nation's Report Card: Fourth-grade students reading aloud: NAEP 2002 special study of oral reading. National Center for Education Statistics. Retrieved from https://files.eric.ed.gov/fulltext/ED488962.pdf

Dowhower SL (1991). Speaking of prosody: Fluency's unattended bedfellow. Theory into Practice, 30, 165–175. Retrieved from https://www.jstor.org/stable/1476878

Foorman B, Petscher Y, & Schatschneider C (2015). Florida Center for Reading Research (FCRR) Reading Assessment: Technical manuals. Tallahassee: Florida Center for Reading Research. Retrieved from http://www.fcrr.org/for-researchers/fra.asp

Fuchs LS, Fuchs D, Hosp MK, & Jenkins JR (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading, 5, 239–256.doi:10.1207/S1532799XSSR0503_3

Gibbons RD, & Hedeker DR (1992). Full-information bi-factor analysis. Psychometrika, 57, 423–436. doi:10.1007/BF02295430

Groen MA, Veenendaal NJ, & Verhoeven L (2019). The role of prosody in reading comprehension: Evidence from poor comprehenders. Journal of Research in Reading, 1, 37–57. DOI:10.1111/1467-9817.12133

Hoover WA, & Gough PB (1990). The simple view of reading. Reading and Writing: An Interdisciplinary Journal, 2, 127–160.

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

Jenkins JR, Fuchs LS, Van den Broek P, Espin C, & Deno SL (2003). Sources of individual differences in reading comprehension and reading fluency. Journal of Educational Psychology, 95, 719–729. doi:10.1037/0022-0663.95.4.719

Kim Y-SG (2015). Developmental, component-based model of reading fluency: An investigation of word-reading fluency, text-reading fluency, and reading comprehension. Reading Research Quarterly, 50, 459–481. doi:10.1002/rrq.107 [PubMed: 26435550]

Kim Y-SG (2020). Hierarchical and dynamic relations of language and cognitive Skills to reading comprehension: Testing the direct and indirect effects model of reading (DIER). Journal of Educational Psychology, 112 (4), 667–684. 10.1037/edu0000407

Kim Y-SG, Quinn J, & Petscher Y (2020). Reading prosody unpacked: A longitudinal investigation of its dimensionality and the relation with word reading and listening comprehension for children in primary grades. Journal of Educational Psychology. 10.1037/edu0000480

Kim Y-S, Park C, & Wagner RK (2014). Is oral/text reading fluency a “bridge” to reading comprehension? Reading and Writing: An Interdisciplinary Journal, 27, 79–99. doi:10.1007/s11145-013-9434-7

Kim Y-S, Petscher Y, Schatschneider C, & Foorman B (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? Journal of Educational Psychology, 102, 652–667. doi:10.1037/a0019643

Kim Y-SG, Petscher Y, & Vorstius C (2019). Unpacking eye movements during oral and silent reading and their relations to reading proficiency in beginning readers. Contemporary Educational Psychology, 58, 102–120. 10.1016/j.cedpsych.2019.03.002

Kim YSG, & Wagner RK (2015). Text (oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from grades 1 to 4. Scientific Studies of Reading, 19(3), 224–242. [PubMed: 25848201]

Kim Y-S, Wagner RK, & Foster E (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. Scientific Studies of Reading, 15, 338–362. doi:10.1080/10888438.2010.493964 [PubMed: 21747658]

Klauda SL, & Guthrie JT (2008). Relationships of three components of reading fluency to reading comprehension. Journal of Educational Psychology, 100, 310–321. DOI:10.1037/0022-0663.100.2.310

Kline R (2016). Principles and practice of structural equation modeling (4th Ed.). New York, NY: Guilford Press.

Kuhn MR, Schwanenflugel PJ, & Meisinger EB (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. Reading Research Quarterly, 45(2), 230–251. DOI:10.1598/RRQ.45.2.4

Kuhn MR, & Stahl SA (2003). Fluency: A review of developmental and remedial practices. Journal of Educational Psychology, 95, 3–21. DOI:10.1037/0022-0663.95.1.3

LaBerge D, & Samuels SJ (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6(2), 293–323. DOI:10.1016/0010-0285(74)90015-2

McDonald RP (1999). Test Theory: A unified approach. Mahwah, NJ: Erlbaum.

Miller J, & Schwanenflugel PJ (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. Reading Research Quarterly, 43(4), 336–354. DOI:10.1598/RRQ.43.4.2 [PubMed: 20072660]

Muthén LK, & Muthén BO (1998–2018). Mplus User's Guide (8th Ed.) Los Angeles, CA: Muthén & Muthén.

Nathan RG, & Stanovich KE (1991). The causes and consequences of differences in reading fluency. Theory Into Practice, 30, 176–184.

National Governors Association Center & Council of Chief State School Officers. (2010). Common core state standards. Washington, D.C.: Author.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00–4769). Washington, DC: Government Printing Office.

Paige D, Rasinski T, Magpuri-Lavell T, & Smith GS (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. Journal of Adolescent and Adult Literacy, 46(2), 123–156. DOI: 10.1177/1086296X14535170

Perfetti CA (1992). The representation problem in reading acquisition. In Ehri LC, Treiman R, & Gough PB (Eds.), Reading acquisition (pp. 145–174). Hillsdale, NJ: Erlbaum.

Perfetti CA, Goldman SR, & Hogaboam TW (1979). Reading skill and the identification of words in discourse context. Memory & Cognition, 7, 273–282. doi:10.3758/BF03197600 [PubMed: 537503]

Raftery AE (1995). Bayesian model selection in social research. Sociological Methodology, 25, 111–163.

Rasinski T (2004). Creating fluent readers. Educational Leadership, 61(6), 46–51. Retrieved from http://educationalleader.com/subtopicintro/read/ASCD/ASCD_364_1.pdf

Ravid D, & Mashraki YE (2007). Prosodic reading, reading comprehension and morphological skills in Hebrew-speaking fourth graders. Journal of Research in Reading, 30(2), 140–156. DOI: 10.1111/j.1467-9817.2007.00340.x

Reise SP (2012). The rediscovery of bifactor measurement models. Multivariate Behavioral Research, 47(5), 667–696. DOI: 10.1080/00273171.2012.715555 [PubMed: 24049214]

Sabatini J, Wang Z, & O'Reilly T (2019). Relating reading comprehension to oral reading performance in the NAEP fourth-grade special study of oral reading. Reading Research Quarterly, 54, 253–271. doi:10.1002/rrq.226

Schreiber PA (1991). Understanding prosody's role in reading acquisition. Theory Into Practice, 30(3), 158–164.

Schwanenflugel PJ, Hamilton AM, Kuhn MR, Wisenbaker J, & Stahl SA (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. Journal of Educational Psychology, 96, 119–129. DOI:10.1037/0022-0663.96.1.119 [PubMed: 19777077]

Seymour PHK, Aro M, & Erskine JM (2003). Foundation literacy acquisition in European orthographies. British Journal of Psychology, 94, 143–174. [PubMed: 12803812]

Shinn M, Good R, Knutson N, Tilly D, & Collins V (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. School Psychology Review, 21, 459–479.

Shrout PE, & Fleiss JL (1979). Intraclass correlations Uses in assessing rater reliability. Psychological Bulletin, 86, 420–428. 10.1037/0033-2909.86.2.420. [PubMed: 18839484]

Torgesen JK, Wagner RK, & Rashotte CA (2012). Test of Word Reading Efficiency–Second Edition (TOWRE-2). Austin, TX: Pro-Ed.

Vadasy PF, & Sanders EA (2008). Repeated reading intervention: Outcomes and interactions with readers' skills and classroom instruction. Journal of Educational Psychology, 100(2), 272–290. DOI: 10.1037/0022-0663.100.2.272

Veenendaal NJ, Groen MA, & Verheoeven L (2014). The role of speech prosody and text reading prosody in children's reading comprehension. British Journal of Educational Psychology, 84, 521–536. DOI:10.1111/bjep.12036

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

Wolf M, & Katzir-Cohen T (2001). Reading fluency and its intervention. Scientific Studies of Reading, 5, 211–239. doi:10.1207/S1532799XSSR0503_2

Woodcock RW, McGrew KS, & Mather N (2001). Woodcock–Johnson III Tests of Achievement. Itasca, IL: Riverside.

Zinbarg RE, Revelle W, Yovel I, & Li W (2005). Cronbach's α, Revelle's β, and McDonald's ω H: Their relations with each other and two alternative conceptualizations of reliability. Psychometrika, 70(1), 123–133. DOI: 10.1007/s11336-003-0974-7

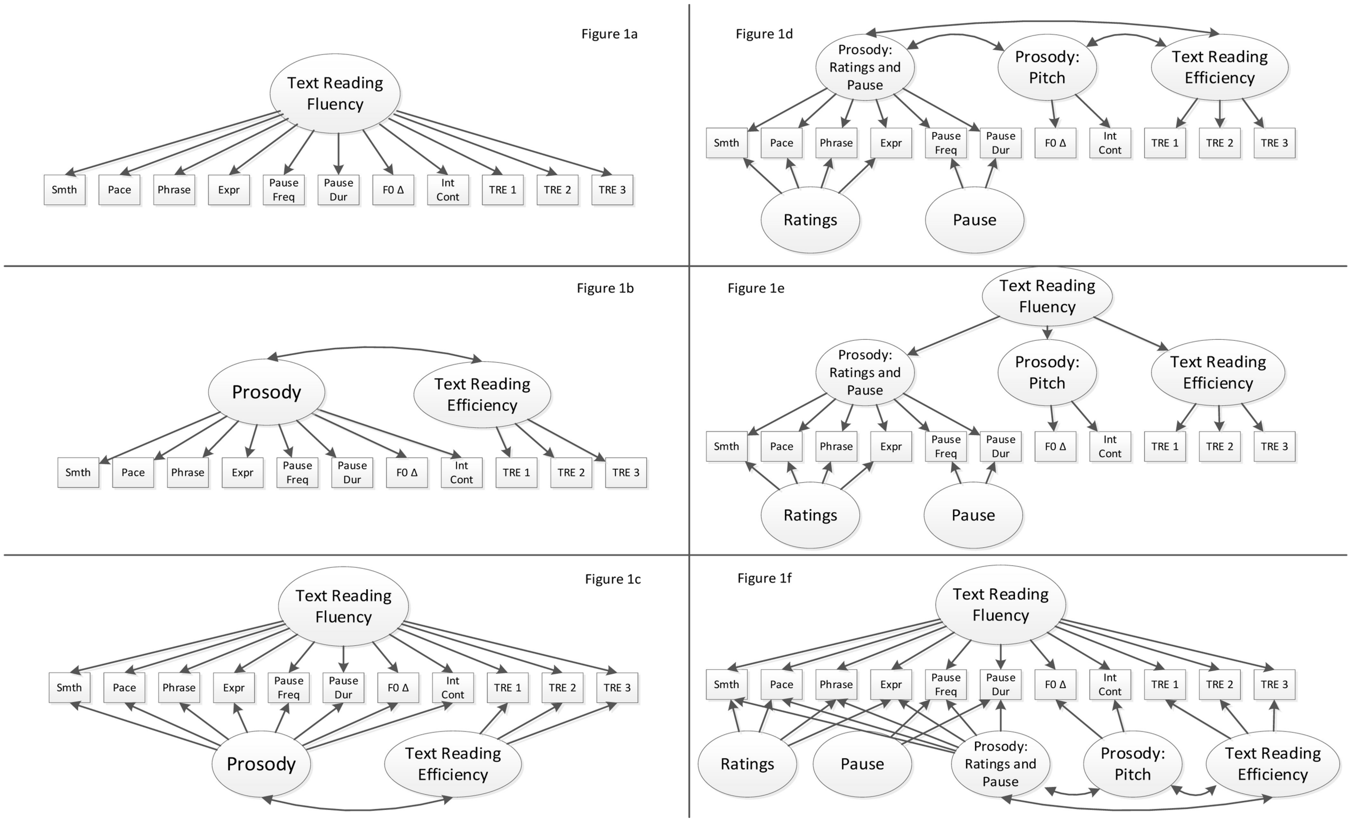

Figure 1.

Dimensionality models fit to the prosody and text reading efficiency variables. *Note*. Smth = Rating scale smoothness; pace = Rating scale pacing; Phrase = Rating scale phrasing; Expr = Rating scale expression & volume; Pause Freq = pause frequencies; Pause Dur = Pause durations; F0 Δ = F0 Change; Int Cont = Intonation Contour; TRE 1- TRE 3 = Text reading efficiency passages 1–3.

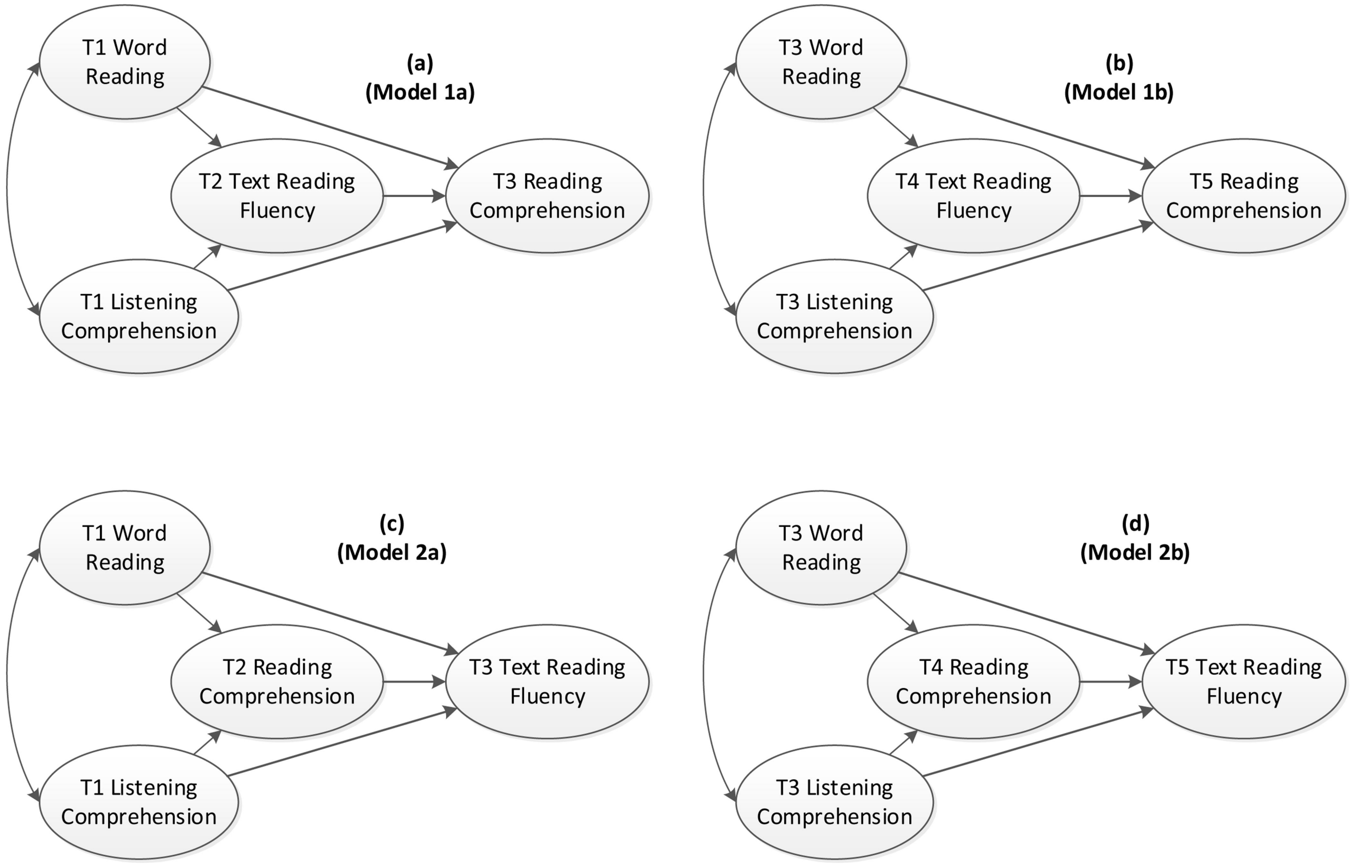

Figure 2.
Mediation models fit to the word reading, listening comprehension, text reading fluency, and reading comprehension data. Models 1a and 1b indicate text reading fluency as the mediator; models 2a and 2b indicate reading comprehension as the mediator. The measurement model portion for word reading (i.e., LWID, WIATWR, SWE), listening comprehension (i.e., OWLS, WJOC), reading comprehension (i.e., WJPC, WIATRC), and text reading fluency (i.e., prosody and text reading efficiency indicators, including specific factors for Ratings, Pause, Prosody: Ratings and Pause, and Prosody: Pitch) was estimated but not included in the figure for ease of interpretation. Baseline measures were also included in the model but excluded from the figure for figure simplicity. T1-T5 = Time 1-Time 5.

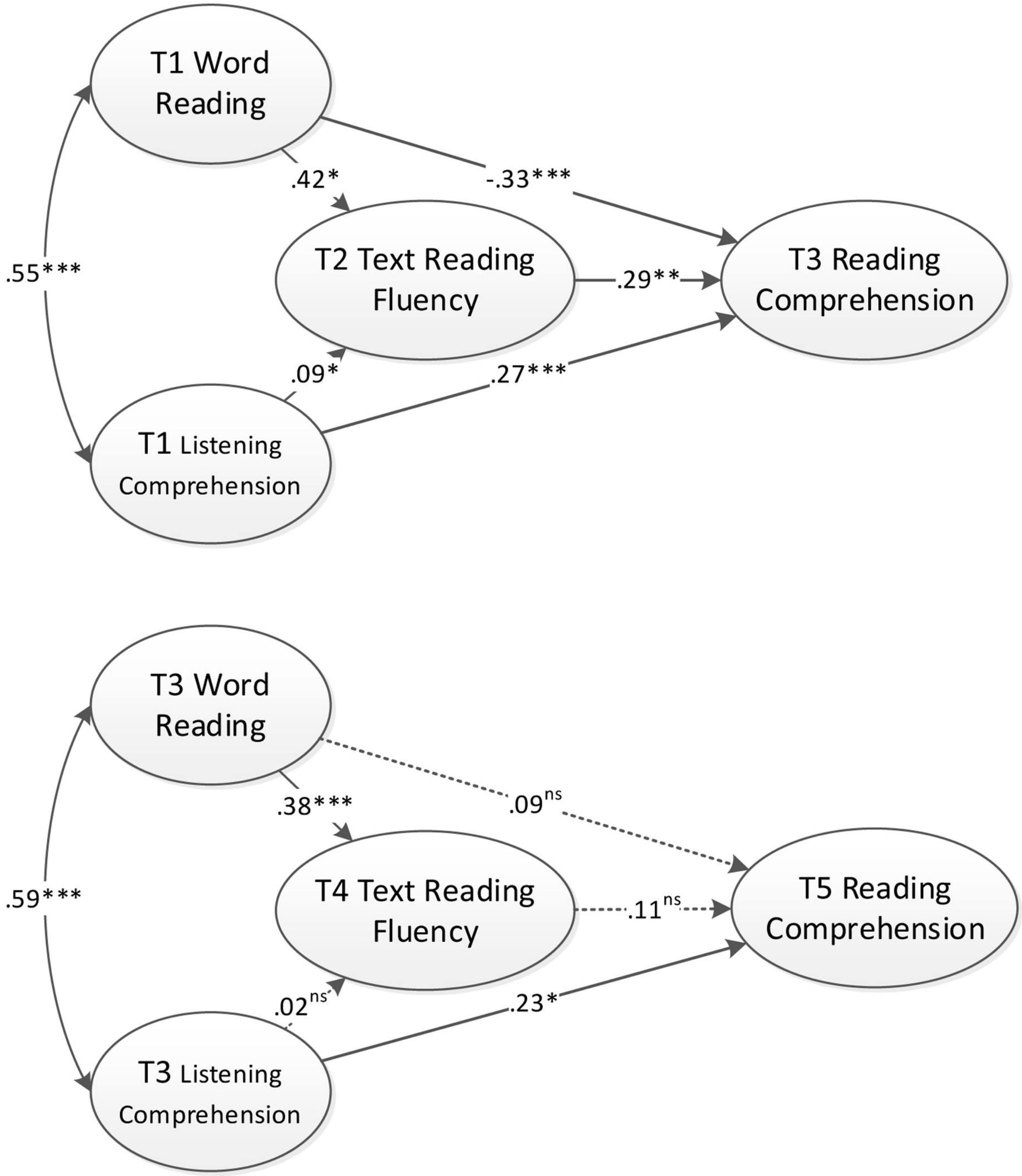

Figure 3.

Standardized path coefficients for the final mediation models for Model 1a (Time 1, Time 2, Time 3) and Model 1b (Time 3, Time 4, Time 5). The full structure of text reading fluency as fit in the dimensionality models and the baseline levels of text reading fluency and reading comprehension were included in each model but were not included in this figure for simplicity. See Figure S2 in the Online Supplemental Materials for full models. T1-T5 = Time 1- Time 5. ^{ns} = not significant at $p < .05$; *** = $p < .001$; ** = $p < .01$; * = $p < .05$

Table 1.

Descriptive statistics for text reading efficiency, reading prosody as measured by spectrographic analysis and the ratings scale, and reading comprehension by time point.

	Time 2					Time 3				
	n	Min	Max	Mean	SD	*n*	Min	Max	Mean	SD
Text Reading Efficiency										
TRE Passage 1	330	4	197	72.21	36.31	313	10	197	82.11	41.36
TRE Passage 2	320	10	175	68.45	33.59	310	11	184	80.04	37.92
TRE Passage 3	327	7	162	73.22	32.41	317	19	199	87.13	37.94
Reading Prosody: Spectrograph										
F0 change	333	−141.4	−19.63	−61.83	21.99	319	−127.9	−1.65	−63.58	21.87
Intonation Contour	337	183.6	331.04	258.19	29.42	320	190.9	321.49	254.61	28.81
Pause Duration	335	0.21	3.06	0.86	0.456	319	0.16	2.08	0.7	0.33
Pause Frequency (Avg.)	335	1	21	9.51	4.86	320	1.33	26	11.54	6.35
Reading Prosody: Ratings Sca										
Smoothness	327	1	4	1.8	0.68	307	1	4	1.96	0.74
Pacing	327	1	4	1.77	0.73	307	1	4	1.98	0.83
Phrasing	327	1	4	1.85	0.73	307	1	4	2.08	0.81
Expression & Volume	328	1	4	2.05	0.73	307	1	4	2.32	0.74
Reading Comprehension										
WIATRC	363	1	36	22.51	7.95	326	0	40	25.57	8.60
WIATRC – SS	363	60	141	104.78	14.50	326	46	149	100.38	14.14
WJPC	363	6	31	21.56	3.90	326	12	35	22.70	4.79
WJPC – SS	363	55	133	106.73	12.45	326	60	126	100.07	11.93
	Time 4					Time 5				
	n	Min	Max	Mean	SD	*n*	Min	Max	Mean	SD
Text Reading Efficiency										
TRE Passage 1	304	16	221	108.49	41.51	301	12	204	108.31	39.34
TRE Passage 2	302	13	188	91.93	38.79	301	23	237	119.54	44.09
TRE Passage 3	300	4	184	89.03	35.67	300	24	240	111.24	39.09
Reading Prosody: Spectrograph										
F0 change	317	−111.6	−9.3	−53.96	21.12	302	−131.5	−15.2	−60.33	23.71
Intonation Contour	318	191.28	320.61	251.15	27.79	302	180.31	316.14	247.2	30.17
Pause Duration	316	0.13	1.84	0.67	0.28	300	0.21	1.62	0.58	0.22
Pause Frequency (Avg.)	317	0	20	9.35	4.79	301	0.33	13.25	5.11	3.18
Reading Prosody: Ratings Scale										
Smoothness	312	1	4	2.34	0.77	276	1	4	2.73	0.84
Pacing	312	1	4	2.39	0.9	276	1	4	2.79	0.82
Phrasing	312	1	4	2.49	0.8	276	1	4	2.93	0.76
Expression & Volume	312	1	4	2.68	0.8	276	1	4	2.97	0.67
Reading Comprehension										
WIATRC	329	9	40	30.29	6.98	311	0	42	28.31	8.44

	Time 2					Time 3				
	n	**Min**	**Max**	**Mean**	**SD**	***n***	**Min**	**Max**	**Mean**	**SD**
WIATRC – SS	329	71	137	104.30	13.32	311	48	145	100.54	14.20
WJPC	329	12	34	24.75	3.38	311	8	37	26.50	4.54
WJPC – SS	329	66	125	101.05	9.63	311	47	127	98.67	11.70

Note. TRE = text reading efficiency; WIATRC = Wechsler Individual Achievement Test – Reading Comprehension; WJPC = Woodcock-Johnson Tests of Achievement – Passage Comprehension; SS = Standard scores. Unless otherwise stated, scores are raw values.

Author Manuscript
Author Manuscript
Author Manuscript

Table 2.

Descriptive statistics for word reading and listening comprehension by time point (n = 350 in Time 1; n = 326 in Time 3)

	Time 1				Time 3			
Measure	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**
Word Reading								
LWID	17	57	32.49	6.75	24	62	41.84	6.60
LWID - SS	68	147	110.99	13.54	73	135	108.48	12.07
WIATWR	1	40	14.10	9.42	5	56	28.00	10.51
WIATWR - SS	60	147	100.15	16.10	63	143	103.27	15.33
SWE	4	67	31.72	14.98	15	80	51.68	13.44
SWE - SS	55	144	100.60	17.59	55	137	103.86	16.57
Listening Comprehension								
WJOC	2	21	13.03	3.64	6	26	16.31	3.55
WJOC – SS	56	134	103.79	13.09	71	140	106.02	12.67
OWLS	40	95	69.80	12.44	46	111	83.81	11.60
OWLS - SS	50	135	100.64	14.08	65	135	106.66	13.06

Note. SS = Standard scores; LWID = Woodcock-Johnson III Letter-Word Identification; WIATWR = Wechsler Individual Achievement Test – Word Reading; SWE = Test of Word Reading Efficiency – Sight Word Efficiency; WJOC = Woodcock-Johnson III– Oral Comprehension; OWLS = Oral and Written Language Scales – Listening Comprehension. Unless otherwise stated, scores are raw values.

Table 3.

Correlation tables for concurrently measured prosody and text reading efficiency variables used for dimensionality models at Time 2 (below diagonal) and Time 3 (above diagonal) in the top panel, and at Time 4 (below diagonal) and Time 5 (above diagonal) in the bottom panel.

	1	2	3	4	5	6	7	8	9	10	11
Time 2 (below diagonal) and Time 3 (above diagonal)											
1. Text reading efficiency 1	--	.92	.92	−.21	.08	−.55	−.86	.82	.82	.78	.76
2. Text reading efficiency 2	.92	--	.90	−.24	.09	−.55	−.87	.79	.82	.78	.77
3. Text reading efficiency 3	.90	.93	--	−.22	.09	−.53	−.84	.80	.81	.78	.77
4. F0 change	−.19	−.17	−.17	--	−.43	.01	.26	−.19	−.23	−.24	−.30
5. Intonation Contour	.02	.03	.02	−.35	--	.00	−.11	.13	.16	.19	.16
6. Pause Duration	−.52	−.54	−.59	.06	.04	--	.50	−.51	−.49	−.47	−.54
7. Pause Frequency	−.80	−.77	−.79	.11	−.02	.46	--	−.76	−.78	−.76	−.74
8. Smoothness	.80	.81	.81	−.17	.05	−.45	−.72	--	.87	.84	.79
9. Pacing	.82	.84	.83	−.16	.12	−.48	−.73	.87	--	.89	.80
10. Phrasing	.81	.82	.83	−.18	.07	−.48	−.75	.84	.88	--	.84
11. Expression & Volume	.73	.72	.76	−.21	.13	−.44	−.70	.78	.80	.84	--
Time 4 (below diagonal) and Time 5 (above diagonal)											
1. Text reading efficiency 1	--	.91	.91	−.24	−.02	−.53	−.86	.79	.83	.79	.72
2. Text reading efficiency 2	.93	--	.90	−.26	.00	−.54	−.86	.78	.84	.78	.72
3. Text reading efficiency 3	.91	.93	--	−.26	.00	−.52	−.83	.76	.81	.76	.70
4. F0 change	−.22	−.25	−.20	--	−.46	.09	.26	−.16	−.24	−.25	−.40
5. Intonation Contour	.10	.10	.10	−.41	--	.05	−.04	−.01	.05	.02	.14
6. Pause Duration	−.54	−.51	−.55	.05	.00	--	.50	−.47	−.52	−.45	−.50
7. Pause Frequency	−.85	−.85	−.83	.26	−.10	.42	--	−.75	−.80	−.80	−.72
8. Smoothness	.82	.76	.75	−.18	.16	−.45	−.74	--	.85	.86	.77
9. Pacing	.86	.82	.81	−.21	.14	−.48	−.79	.89	--	.86	.80
10. Phrasing	.84	.80	.78	−.25	.10	−.46	−.78	.87	.89	--	.81
11. Expression & Volume	.76	.72	.70	−.33	.19	−.44	−.71	.82	.81	.84	--

Note. Correlation values inclusive of $-.11 \leq r \leq .11$ are not significant at $p < .05$.