CHILD DEVELOPMENT

Child Development, xxxx 2017, Volume 00, Number 0, Pages 1–18

Unpicking the Developmental Relationship Between Oral Language Skills and Reading Comprehension: It's Simple, But Complex

Arne Lervåg
University of Oslo

Charles Hulme
University of Oxford

Monica Melby-Lervåg
University of Oslo

Listening comprehension and word decoding are the two major determinants of the development of reading comprehension. The relative importance of different language skills for the development of listening and reading comprehension remains unclear. In this 5-year longitudinal study, starting at age 7.5 years (n = 198), it was found that the shared variance between vocabulary, grammar, verbal working memory, and inference skills was a powerful longitudinal predictor of variations in both listening and reading comprehension. In line with the simple view of reading, listening comprehension, and word decoding, together with their interaction and curvilinear effects, explains almost all (96%) variation in early reading comprehension skills. Additionally, listening comprehension was a predictor of both the early and later growth of reading comprehension skills.

The ability to read text with understanding is one of the core aims of primary school education, and adequate reading comprehension skills are essential for educational success and adult well-being. It is well established that the development of reading comprehension depends critically on word decoding and listening comprehension (the simple view of reading; Hoover & Gough, 1990; Gough & Tunmer, 1986). However, many questions remain about the relative importance of different oral language skills (e.g., vocabulary, grammatical, and inferential skills) as influences on the development of listening and reading comprehension. There is also a lack of consensus about the exact form of the relationships between word reading and listening comprehension as determinants of reading comprehension. Here we answer these critical questions using data from a large-scale longitudinal study.

The Simple View of Reading

There are several different models that have been used as frameworks for understanding how reading comprehension develops (Cromley & Azevedo, 2007; Kintsch, 1988; McNamara & Kintsch, 1996; Perfetti & Stafura, 2014). However, for elementary school children, by far the most commonly cited theoretical framework is the simple view of reading (Gough & Tunmer, 1986). In this view, understanding written text is the product of decoding and listening comprehension. Decoding refers to the ability to convert print into sound and to read fluently (see NICHD Early Child Care Research Network, 2005). The simple view implies that when decoding skills are poor, they will place important constraints on reading comprehension. In contrast, when decoding skills are stronger, listening comprehension becomes a more important influence on reading comprehension. In the last 30 years, the simple view of reading has been used in a number of studies across different languages, particularly in studies of children in early elementary school but also in some studies of later elementary school children (see García & Cain, 2014 for a meta-analysis).

The basic tenets of the simple view of reading are supported by a large body of evidence. According to a recent meta-analysis (García & Cain, 2014), there is a strong concurrent correlation between decoding and

This research was funded by the Department of Education, University of Oslo and by a grant from the Research Council of Norway's UTD2020 programme (Grant 203335). We thank the research assistants for their help with data collection, and the teachers, schools, children, and parents for their participation.

Correspondence concerning this article should be addressed to Arne Lervåg, Department of Education, University of Oslo, P.O. Box 1092 Blindern, NO-0317 Oslo, Norway. Electronic mail may be sent to a.o.lervag@iped.uio.no.

0009-3920/2017/xxxx-xxxx
DOI: 10.1111/cdev.12861

reading comprehension (r = .74), though, as predicted, this correlation becomes weaker in older age groups where the correlation between listening comprehension and reading comprehension becomes stronger (see also Chen & Vellutino, 1997; Foorman, Koon, Petscher, Mitchell, & Truckenmiller, 2015). A critical limitation of this evidence is that it comes from concurrent studies. Evidence from longitudinal studies is needed to provide evidence of putative causal effects.

Longitudinal studies of reading comprehension typically concentrate on the early stages of learning to read, typically up to third grade (Aarnoutse, van Leeuwe, & Verhoeven, 2005; Kendeou, Van den Broek, White, & Lynch, 2009; Näslund, 1990; NICHD Early Child Care Research Network, 2005; Roth, Speece, Cooper, & De la Paz, 1996); studies that follow children over longer periods of time (i.e., to fourth grade or later) are scarce (Geva & Farnia, 2012; Limbird, Maluch, Rjosk, Stanat, & Merkens, 2014; Storch & Whitehurst, 2002; Verhoeven & van Leeuwe, 2012). The main finding from these studies is that reading comprehension can be largely predicted from listening comprehension and word decoding. As age increases, the role of word decoding as a predictor of reading comprehension (after reading comprehension at an earlier time point has been taken into account) decreases and the role of listening comprehension increases (Verhoeven & van Leeuwe, 2012; see also Geva & Farnia, 2012; Storch & Whitehurst, 2002).

These studies all use autoregressive models that essentially assess whether the rank order of children changes over time, but they do not model differences in relative rates of development among children. One exception is the study by Quinn, Wagner, Petscher, and Lopez (2015), which used a latent change score model. This study provides support for a causal influence of vocabulary on reading comprehension because previous levels of vocabulary knowledge acted as crucial drivers of reading comprehension growth (Quinn et al., 2015). Another limitation of most longitudinal studies is that few have used latent variables with multiple indicators of each construct to control for measurement error (e.g., de Jong & van der Leij, 2002; Storch & Whitehurst, 2002).

Unresolved Issues in the Development of Reading Comprehension

Although the main elements of the simple view of reading have been strongly supported by previous research, there remain several unresolved theoretical issues concerning the nature and form of influences on the development of reading comprehension.

The Components of Listening Comprehension and Their Relationship With Reading Comprehension

Although listening comprehension appears to be a crucial influence on reading comprehension (Gough & Tunmer, 1986), the nature of the language skills that provide the foundations for listening comprehension need to be clarified. Evidence suggests that vocabulary knowledge is one critical influence on listening comprehension (Clarke, Snowling, Trulove, & Hulme, 2010; Kim, 2015, 2016; Lervåg & Aukrust, 2010; Protopapas, Mouzaki, Sideridis, Kotsolakou, & Simos, 2013; see also Sénéchal, Ouellette, & Rodney, 2006).

Few studies, however, have examined the possible role of other language-related skills as predictors of listening comprehension. It has been suggested that variations in verbal working memory capacity may place constraints on listening comprehension. For example, a study by Florit, Roch, and Levorato (2011) found that preschool children's working memory was a significant concurrent predictor of listening comprehension after the effects of vocabulary were controlled (see also Dufva, Niemi, & Voeten, 2001; Kim, 2016). However, the opposite has also been suggested, that is, that performance on verbal working memory tasks merely reflect variations in language skills (MacDonald & Christiansen, 2002; see also Klem et al., 2015; Melby-Lervåg et al., 2012).

One other cognitive skill that is clearly related to language skills is the ability to draw inferences from spoken texts. It has been suggested that inferential skills are critical for the development of listening comprehension, and Lepola, Lynch, Laakkonen, Silvén, and Niemi (2012) found that these skills were a unique predictor of later listening comprehension beyond vocabulary and prior listening comprehension skills.

It has also been suggested that syntax (understanding the rules governing how words are combined to convey different meanings) is a critical determinant of listening comprehension. Two concurrent studies showed that listening comprehension was directly predicted by inference skills in addition to grammatical knowledge and verbal working memory (Kim, 2015, 2016).

Influences on Reading Comprehension Beyond Listening Comprehension?

According to the simple view of reading (Gough & Tunmer, 1986), once a text has been decoded, the only limit on comprehension is variations in listening comprehension. This is a simple and radical proposal. However, others have suggested that other skills, including the language skills underlying listening comprehension, may have direct effects on reading comprehension which are not fully mediated by listening comprehension (Geva & Farnia, 2012; Joshi & Aaron, 2000; Kirby & Savage, 2008; Silva & Cain, 2015). These authors have argued "for a slightly less simple view of reading" (Kirby & Savage, 2008, p. 75).

For example, it has been suggested that verbal working memory is crucial because reading comprehension requires the ability to process and store information concurrently (Daneman & Carpenter, 1980; Just & Carpenter, 1992). Some cross-sectional studies (Cain, Oakhill, & Bryant, 2004; Christopher et al., 2012) and one longitudinal study (Seigneuric & Ehrlich, 2005) have provided support for this claim, finding that verbal working memory uniquely explains variations in reading comprehension after controlling for the effects of vocabulary, decoding, and earlier reading comprehension. However, another cross-sectional study failed to find any predictive relationship between verbal working memory and reading comprehension after controlling for decoding, listening comprehension, and vocabulary (Cutting & Scarborough, 2006).

It has also been suggested that inferential skills are critical for the development of reading comprehension and that it has a direct influence on reading comprehension in addition to its effect on listening comprehension (Cromley & Azevedo, 2007; Kintsch, 1988; Oakhill & Cain, 2000; Yuill & Oakhill, 1991; see also Perfetti & Stafura, 2014). Consistent with this, Oakhill and Cain (2012) demonstrated that inference skills predicted reading comprehension after controlling for prior levels of reading comprehension ability.

Finally, it has been suggested that syntax is not only vital for listening comprehension but also has a direct influence on reading comprehension as it, together with word meaning, constitutes a lexicon that is critical for comprehension processes (see figure 1 in Perfetti & Stafura, 2014). However, Kim (2015) found that the effects of syntax were mediated through listening comprehension and that it had no additional direct effect on reading comprehension.

Moderation of the Relationship Between Listening Comprehension and Reading Comprehension

The simple view of reading (Gough & Tunmer, 1986) claims that reading comprehension reflects a multiplicative relationship between decoding and listening comprehension ($R = D \times C$; Reading Comprehension = Decoding × Listening Comprehension) rather than a simple additive one. This implies that the associations between reading comprehension, listening comprehension, and decoding will change during the course of development. Early in development, decoding skills will vary widely and provide powerful constraints on reading comprehension. Conversely, among older children who have proficient decoding skills, reading comprehension will be more heavily influenced by listening comprehension skills. Studies that have examined the possible multiplicative effect of decoding and listening comprehension as determinants of reading comprehension have produced highly inconsistent results. In a seminal study, Hoover and Gough (1990) examined this relationship in a sample of second-language learners in kindergarten through fourth grade. The best fit to the data was obtained with a regression model that included a product term in addition to the linear effects of decoding and listening comprehension. In contrast, Chen and Vellutino (1997) examined the simple view of reading in a concurrent study with samples of monolingual English children in Grades 2, 3, 6, and 7, and failed to find any evidence for the interaction.

The Current Study

An inconsistent pattern emerges from prior studies of the relationship between decoding, component language skills, listening comprehension, and reading comprehension skills. Some of these inconsistencies may reflect a failure to take account of measurement error (which may serve to distort the pattern of predictive relationships present if measures have different reliabilities). Measurement error will also serve to reduce the power to detect interactive effects between decoding and listening comprehension ($R = D \times C$) because when the reliability of measures is less than perfect, product terms are inherently less reliable than the simple terms they are derived from. One other important methodological point is that it is critical to assess the extent to which predictors show linear (rather than curvilinear) relationships with reading comprehension (see Ganzach, 1997).

The current study provides a clearer picture by using a large set of predictors that allow us to examine the underlying structure of listening comprehension and how it relates to other language-related skills, such as vocabulary, verbal working memory, inference skills, and grammar. We use latent variables to control for measurement error. We test the following hypotheses:

1. Individual differences in listening comprehension will primarily reflect variations in a latent language factor that can be measured by diverse measures of component language skills (vocabulary, grammatical [syntactic and morphologic skills], verbal working memory, and inference skills).
2. There will be an interactive effect of decoding and listening comprehension on reading comprehension (R = D × C) such that variations in decoding will place stronger constraints on reading comprehension earlier in development.
3. As listening comprehension is believed to become more important for reading comprehension in older children, we expect that listening comprehension will also predict variations in the growth of reading comprehension skills.
4. Finally, we test "a less simple view of reading" by examining whether the specific factors of vocabulary, grammar, verbal working memory, and inference skills predict the development of reading comprehension after accounting for the effects of listening comprehension.

Method

Participants

One hundred and ninety-eight Norwegian second-grade children (93 girls and 105 boys) were recruited 4 months after formal reading instruction had started (average age 7 years 6 months). All spoke Norwegian as their first language (L1). None of the children had diagnosed developmental disabilities or sensory impairments at the beginning of the study. Informed consent for the children to participate was obtained from their parents. In Norway, children begin school in August of the year that they turn 6 years of age. Formal literacy instruction started in the first grade for all the children in the sample. The children were recruited from schools located in working-class and middle-class areas. The sample attrition was 0.5%, 7.6%, 8.1%, 11.6%, and 19.2% at Times 2, 3, 4, 5, and 6, respectively. The main cause of attrition was children moving out of a school's catchment area.

Design and Procedure

The children were tested on six occasions over a period of 5 years (December 2006–January 2012), starting in the middle of Grade 2 (age 7.5 years) and ending in the middle of Grade 7. From the middle of Grade 2 until the end of Grade 3, children were tested on four occasions at 6-month intervals. Subsequently, they were tested in the middle of Grade 6 (i.e., after a 2.5-year interval) and the middle of Grade 7 (i.e., after a 1-year interval). All testing was performed individually in school, and the tests were given in a fixed order to all participants. In order to minimize possible fatigue effects for the later measures, and to reduce the stress on the children, the test battery was divided into three and administered on different days. Test administrators were instructed to give children breaks if there were any signs of fatigue.

Tests and Materials

At Time 1, nine tests were used to measure language comprehension, two to measure word decoding, and one to measure reading comprehension. The reading comprehension test was readministered at Times 2–6.

Reading comprehension was measured using a Norwegian translation of the Neale Analysis of Reading Ability, 2nd ed. (NARA–II) Form A (Neale, 1997). The test consists of six stories of increasing difficulty. The children were asked to read each story aloud and answer four questions about the first story and eight questions about each of the other five stories. The test administrator asked all questions, and the test was discontinued after the child reached the number of decoding errors specified in the manual. The test involves open-ended questions and narrative texts that draw more heavily on the comprehension component of reading comprehension than tests with multiple choice or cloze procedures (Bowyer-Crane & Snowling, 2005; Cain & Oakhill, 2006; Keenan, Betjemann, & Olson, 2008).

Word decoding was measured using a Norwegian translation of the Test of Word Reading Efficiency (TOWRE) Forms A and B (Torgesen, Wagner, & Rashotte, 1999). The children read as many words as they could in 45 s from a list of 104 words.

Vocabulary was measured using the vocabulary test from the Norwegian adaption of the Wechsler Intelligence Scale for Children, 3rd ed. (WISC–III; Wechsler, 2003) and a Norwegian translation of the first 144 words of the Peabody Picture Vocabulary Test, 3rd ed. (PPVT–III—Form A; Dunn & Dunn, 1997). The tests were administered according to the test manual, except that all children started at Set 3 (ages 6–7) on the PPVT.

Grammatical skills were measured using the Norwegian adaption of the Test for Reception of Grammar, Version 2 (TROG-2; syntactic skills; Bishop, 2009) and the grammatic closure test (morpheme generation) from the Illinois Test of Psycholinguistic Abilities (ITPA; Gjessing & Nygaard, 1995).

Verbal working memory was measured using a Norwegian translation of the listening recall subtest from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). In this test, the child had to listen to sentences read aloud by the administrator, judge whether each sentence was true or false, and then repeat the last word in each sentence in the correct order.

Listening comprehension was measured using a Norwegian translation of the oral comprehension test from the Woodcock–Johnson III battery (Woodcock, McGrew, & Mather, 2001) and a Norwegian translation of NARA–II Form B (Neale, 1997). The Oral Comprehension test from the Woodcock–Johnson III battery is a cloze test in which the child is required to complete a sentence or paragraph with an appropriate word.

In our listening comprehension version of the NARA–II test, we administered six stories of increasing difficulty. The test administrator read each story aloud and the child had to answer four questions about the first story and eight questions about each of the other five stories. To emphasize comprehension rather than memory, the test administrator stopped after reading approximately half of each of the stories with eight questions and asked the first four questions. After the child had answered those questions, the tester read the rest of the story and then administered the last four questions. The test stopped after a score of zero was obtained for two consecutive stories.

Inference skills were measured using two experimental tasks modeled after those used by Cain, Oakhill, and Elbro (2003). In these tasks, the administrator read aloud a sentence containing a nonword. The child was then asked if she or he could tell what the nonword meant. On the first presentation, there was not enough contextual information for the child to answer the question; it was merely an introduction to the nonword. The child was then asked to work out the meaning of the nonword by listening to the rest of the story. After reading the rest of the story, which contained sufficient contextual information to provide a reasonable explanation of the meaning of the nonword, the administrator again asked the child what the nonword meant. The answer was given a score of 0, 1, or 2 according to specific criteria for each nonword. A child scored 2 points for a full definition and 1 point for a relatively vague definition. The instructions, an example of one of the stories, and examples of the scoring criteria are presented in Data S1. Two tests were administered on different days, and each contained eight stories. The first six of these stories contained one nonword, and the last two stories contained two nonwords. The maximum possible score on each of these tests was 20.

Results

Descriptive statistics for all measures are presented in Table 1. All measures had relatively high reliabilities at all time points and none of the children reached ceiling on any of the measures. The correlations between measures at all time points are shown in Table S1. All further analyses were performed using full information maximum likelihood estimators with robust (clustered) standard errors as implemented in Mplus version 7.4 (Muthén & Muthén, 1998–2015).

Analyses

In order to test our hypotheses, we used structural equation models with latent variables. In these analyses we compared how well our hypothesized models fitted the data and if comparable (nested) models differed significantly from each other. A good fit tells us that the model is plausible given the data. Following the recommendations of Hu and Bentler (1999), a good model fit was indicated by a root mean square error approximation (RMSEA) < .06 combined with a standardized root mean square residual (SRMSR) < .06 or a comparative fit index (CFI)/Tucker–Lewis index (TLI) of above .95–.96 in combination with an SRMSR below .08.

In addition, a nonsignificant chi-square value for a model indicates that there is no significant difference between the model (the model implied covariance matrix) and the data (the estimated covariance matrix). In order to test the difference between comparable models, we used chi-square difference tests.

Growth in Reading Comprehension Skills

Figure 1 shows the observed individual growth curves between the middle of Grade 2 and the middle of Grade 7. To fit these growth curves, we estimated a piecewise growth model in which the children's growth in reading comprehension skills

Table 1
Number of Participants, Means, Standard Deviations, Range, Skewness, Kurtosis, and Cronbach's Alpha for All Variables at All Time Points

	n	*M*	*SD*	Range	Skewness	Kurtosis	α
Reading comprehension							
T1 NARA	198	9.71	5.93	0–28	.695	0.082	.883
T2 NARA	197	13.94	6.85	0–34	.421	−0.019	.892
T3 NARA	183	18.08	7.06	0–35	.047	−0.334	.871
T4 NARA	182	22.36	7.41	0–40	−.399	0.021	.889
T5 NARA	175	28.63	7.34	0–43	−.615	0.579	.876
T6 NARA	160	30.89	6.61	4–43	−1.059	2.065	.866
Vocabulary							
T1 WISC–III vocabulary	198	18.55	4.19	6–30	.157	−0.023	.721
T1 PPVT–III	198	99.16	14.25	56–137	−.474	−0.302	.930
Grammar							
T1 TROG–2 syntactic skills	198	13.54	3.38	3–20	−.821	1.904	.742
T1 ITPA grammatic closure	198	22.17	4.30	10–33	−.209	−0.167	.757
Working memory							
T1 WMTB-C listening recall	198	9.69	3.26	0–17	−.217	−0.183	.809
Inference skills							
T1 Inference skills A	198	12.61	5.03	0–22	−.555	−0.317	.754
T1 Inference skills B	198	10.29	4.43	0–21	−.232	−0.206	.705
Listening comprehension							
T1 NARA listening comprehension	198	15.12	5.38	3–32	.125	0.117	.820
T1 W–J listening comprehension	198	15.20	3.44	6–25	.277	0.144	.702
Word decoding							
T1 TOWRE A	198	27.67	15.05	0–74	.535	0.191	.974[a]
T1 TOWRE B	198	26.07	17.06	0–75	.720	0.096	.974[a]

NARA = Neale Analysis of Reading Ability; WISC–III = Wechsler Intelligence Scale for Children, 3rd ed.; PPVT–III = Peabody Picture Vocabulary Test, 3rd ed.; TROG = Test for Reception of Grammar, Version 2; ITPA = Illinois Test of Psycholinguistic Abilities; WMTB-C = Working Memory Test Battery for Children; W–C = Woodcock–Johnson; TOWRE = Test of Word Reading Efficiency.
[a]Correlation between Form A and Form B.

was represented by three latent growth constructs. The first construct reflected the score at the first time point in the middle of second grade (initial status), the second reflected the early linear growth rate per year from the middle of Grade 2 to the end of Grade 3 (early growth), and the third reflected the later linear growth rate per year from the end of Grade 3 to the middle of Grade 7 (later growth). As these growth constructs are latent variables, measurement error will not influence their relationship with other latent variables. This model provides an excellent fit to the data: $\chi^2 = 7.162$ (12), $p = .847$; CFI = 1.00; TLI = 1.01; RMSEA = .000, 90% CI [.000, .041]; SRMSR = .032. Figure 2 shows a graphical description of the model together with the means of the growth constructs, the correlations between them, and the standardized residual variances of the observed variables. The model shows that on average, the children answered 8.36 more questions correctly each year between the middle of second grade and the end of third grade (early growth), and 2.49 more questions per year correctly between the middle of third grade and the middle of seventh grade (later growth).

The Structure of Listening Comprehension and Word Decoding

Before predicting the growth of reading comprehension skills, we used confirmatory factor analyses (CFA) to examine the structure of the predictor variables. This allowed us to test if it was possible to capture all the variance from the language, inference, and verbal working memory variables in a common language factor. The model fit indices for the different models are listed in Table 2. First, we estimated a six-factor CFA model in which WISC vocabulary and PPVT reflected a vocabulary factor, NARA and Woodcock–Johnson listening comprehension reflected a listening comprehension factor, the two tests of inference skills reflected an inference factor, TROG and ITPA grammatical closure reflected a grammatical skills factor, and the two TOWRE tests reflected a word decoding factor. To

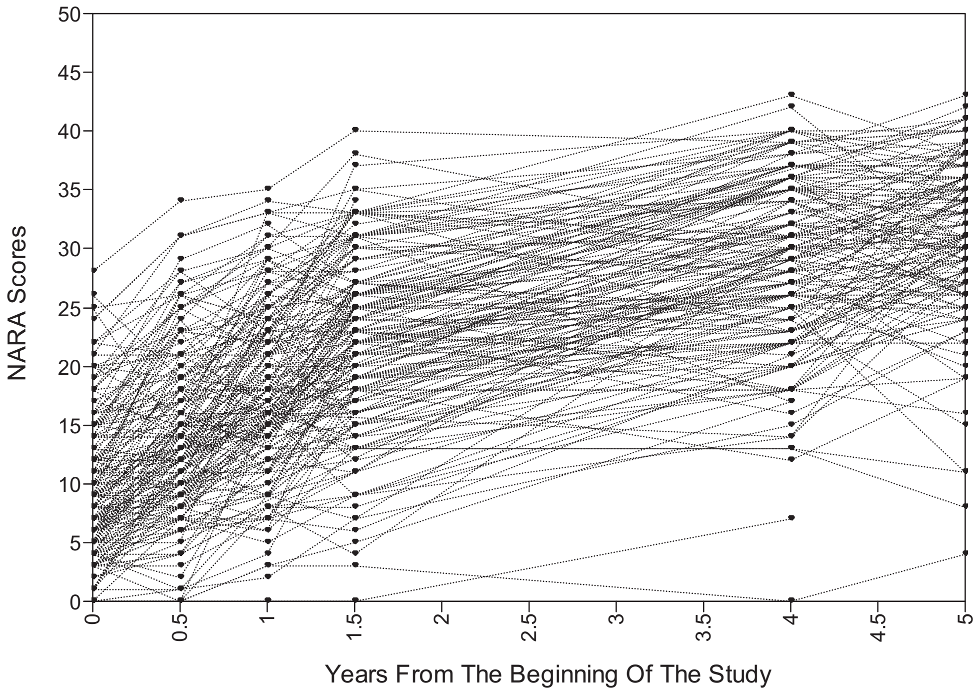

Figure 1. Individual growth curves for reading comprehension from the middle of second grade to the middle of seventh grade.

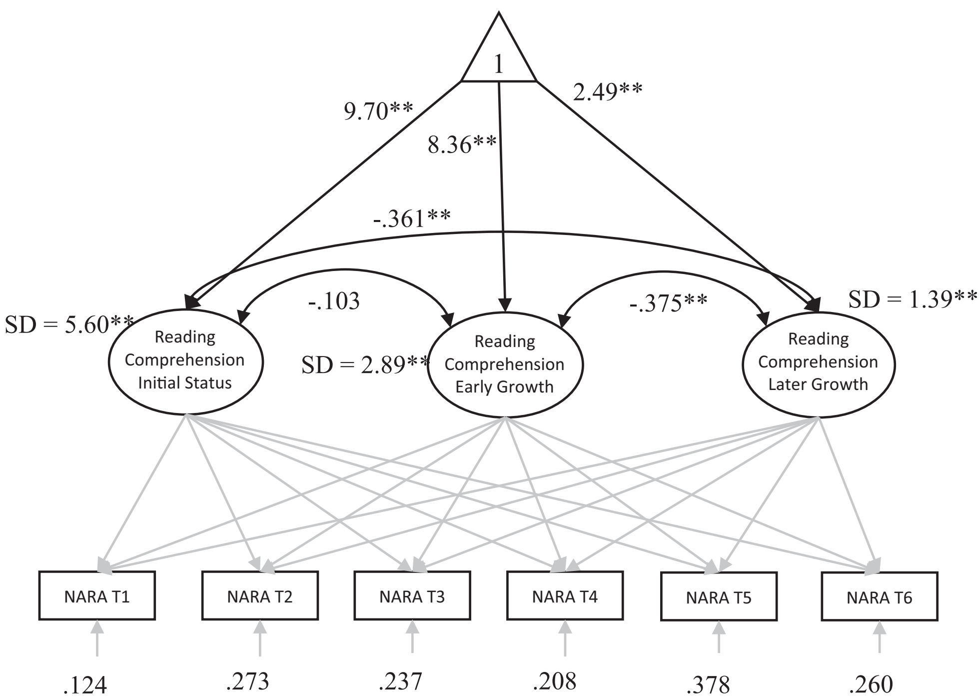

Figure 2. The unconditional growth model of reading comprehension skills from the middle of second grade to the middle of seventh grade. The means are unstandardized, but the correlations shown between the growth constructs and the residuals of the observed variables are standardized. ** = $p < .01$.

create a latent variable for verbal working memory, and thereby estimate verbal working memory without measurement error, we used item parceling and divided the test into odd- versus even-numbered items. As can be seen from Table 2, this six-factor model provided an excellent fit to the data, which

Table 2
Fit Indices for the Confirmatory Factor Analyses Models of the Language, Listening Comprehension, and Word Decoding Tasks

	$\chi^2(df)$, *p*	$\Delta\chi^2(df)$, *p*	RMSEA [90% CI]	CFI	TIL	SRMSR
1. Six-factor model	39.704 (42), .572		000 [.000, 044]	1.00	1.00	.030
2. Second-order model	59.458 (50), .169	20.810 (8) .008	.031 [.000, .058]	.992	.990	.045
3. Modified second-order model (Figure 3)	52.516 (49), .339	13.161 (7), .068	.019 [.000, .051]	.997	.996	.042

RMSEA = root mean square error approximation; CFI = comparative fit index; TLI = Tucker–Lewis index; SRMSR = standardized root mean square residual.

demonstrates that the different language measures cluster together well in their hypothesized categories (vocabulary, grammar, inference, verbal working memory, listening comprehension, and word decoding).

Next, we estimated a model in which the four first-order factors of vocabulary, grammatical skills, verbal working memory, and inference skills reflected a second-order language factor. This second-order factor reflects the common variance shared by the four first-order constructs. This was important because we wanted to determine whether the unique parts of these four factors could predict the development of listening comprehension and reading comprehension after what is common to all of them (i.e., the language factor) is controlled. This method of partitioning the language part of these four constructs into a common language factor made it possible for us to use the unique error-free variance of the first-order factors (the residuals) as predictors. This model fitted the data very well, but the fit was significantly worse than that of the six-factor first-order model indicating that there was more than one general language dimension in the data (see Table 2).

The modification indices for this model revealed that verbal working memory and inference skills shared variance in common (a dimension) that was not accounted for by the common language factor. When allowing the residuals of these two factors to correlate, there was no significant difference between this second-order model and the six-factor first-order model, scaled $\Delta\chi^2 = 13.161(7)$, $p = .068$. We judged the correlation between the unique parts of verbal working memory and inference skills to be justifiable for two reasons: first, they both involve inference skills; second, because our main purpose was to determine whether any of the five first-order factors could predict reading comprehension beyond the common variance for all of them, the correlation between the residuals of verbal working memory and inference skills did not matter. This final CFA model provided an excellent fit to the data (see Table 2).

Figure 3 shows the structural equation model where listening comprehension was regressed on the common language factor (same model fit in the last CFA above). As Figure 3 shows, the standardized path from language to listening comprehension was strong (.95), and language explained 90% of the variance in listening comprehension. Thus, the common language factor is almost isomorphic with the latent listening comprehension factor. The factor loadings of the language factor were high for all of the four first-order constructs except verbal working memory, which had a moderate factor loading. The correlation between the language factor and word decoding was weak, and there was no correlation between listening comprehension and word decoding when we controlled for language.

To test whether the first-order factors were associated with listening comprehension after the effects of the language factor were accounted for, we regressed listening comprehension directly on the residuals (residual factors) one at a time. These paths were not significant for any of the residuals (vocabulary: $p = .805$; grammatical skills: $p = .428$; verbal working memory: $p = .067$; inference skills: $p = .873$), showing that only what they have in common is related to listening comprehension.

Additive, Product, or Nonlinear Effects of Listening Comprehension and Word Decoding on Reading Comprehension

We estimated two models to assess the role of the interaction term (Listening Comprehension × Decoding) in predicting reading comprehension. First, we ran an additive model and then a product model with interactions and curvilinear effects included. In these models, later growth constructs were regressed on the earlier growth construct to avoid confounding the relationship between the earlier growth constructs and the two

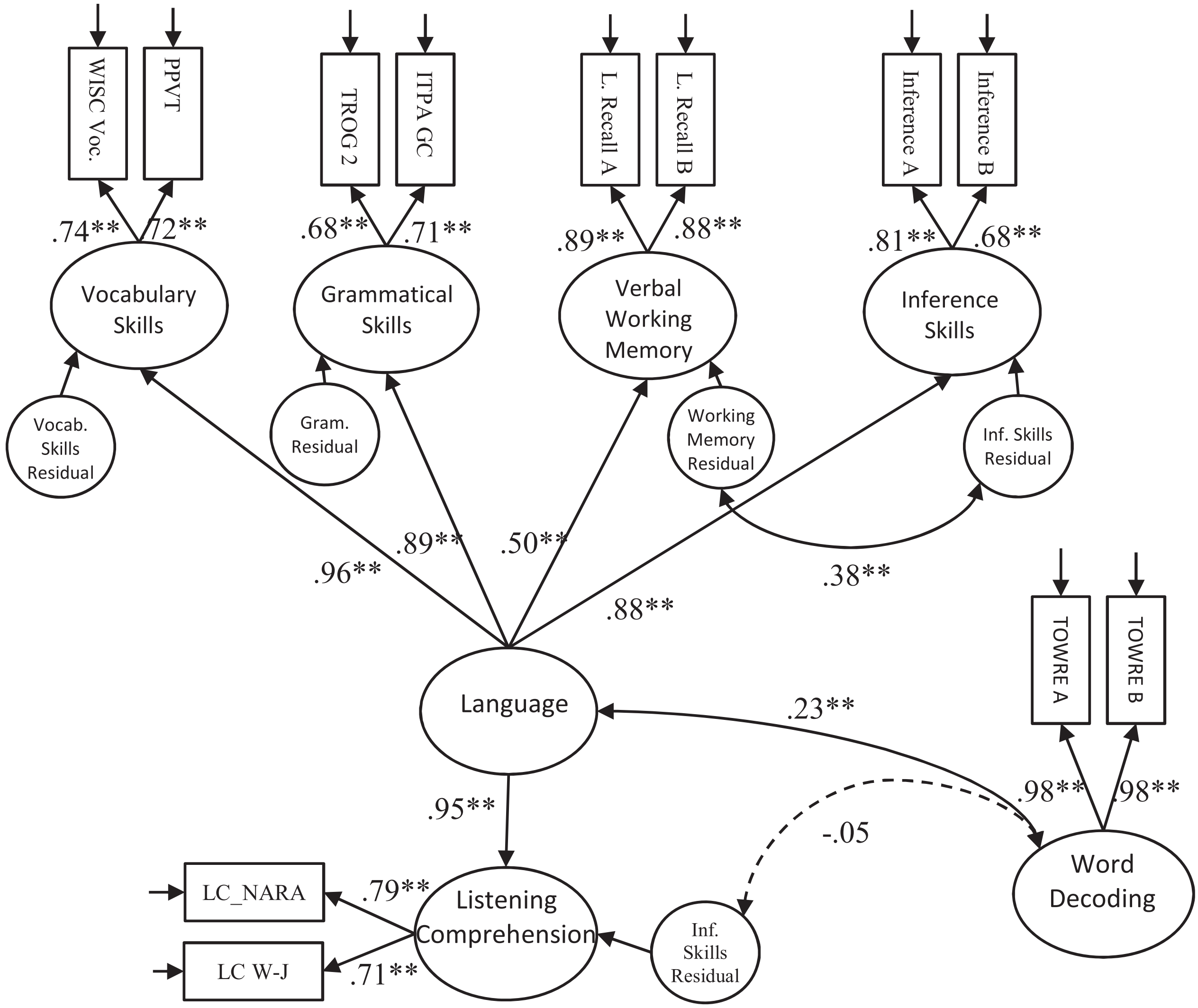

Figure 3. The structure of listening comprehension and word decoding: listening comprehension regressed on the common language skills of vocabulary, grammar, working memory, and inference skills. Parameters shown are standardized. ** = $p < .01$.

predictors of listening comprehension and word decoding.

The left portion of Table 3 shows the results of the additive model. Both listening comprehension and word decoding were strong predictors of initial status in reading comprehension skills. Seventy-seven percent of the variance in the initial status of reading comprehension was explained by this model. Furthermore, both listening comprehension and word decoding predicted early growth in reading comprehension after initial status was controlled; in addition, listening comprehension, but not word decoding, predicted later growth. The negative associations between initial status and early/later growth and between early growth and later growth indicate that children who start out with relatively poor reading comprehension tend to catch up with their peers. This model provided an excellent fit to the data: $\chi^2 = 134.89$ (128), $p = .321$; CFI = .997; TLI = .996; RMSEA = .016, 90% CI [.000, .039]; SRMSR = .045.

The right portion of Table 3 shows the results of the product model, with both interactions and curvilinear effects included. Word decoding moderated the path from listening comprehension to the initial status of reading comprehension (significant interaction). For good decoders, the relationship between listening comprehension and the initial status of reading comprehension was stronger than for poor decoders. Furthermore, the relationships between the initial status of reading comprehension and the two predictors were nonlinear (as indicated

Table 3

Growth in Reading Comprehension Predicted From Linguistic Comprehension and Word Decoding.

	Additive model			Product model with interaction and curvilinear effect		
	Initial status β [95% CI]	Early growth β [95% CI]	Later growth β [95% CI]	Initial status β [95% CI]	Early growth β [95% CI]	Later growth β [95% CI]
Initial status	—	**−1.02** [−1.54, .506]	**−1.21** [−1.88, .533]	—	**−2.03** [−.342, .640]	−1.54 [−3.42, .349]
Early growth	—	—	**−.629** [−1.10, .156]	—	—	**−.659** [−1.16, .157]
Linguistic comprehension	**.700** [.630, .769]	**.944** [.510, 1.38]	**.863** [.288, 1.44]	**.679** [.609, .750]	**1.49** [.464, .209]	1.09 [−.260, 2.45]
Word decoding	**.397** [.312, .483]	**.424** [.142, .707]	.182 [−.147, .510]	**.485** [.396, .574]	**1.01** [.347, 1.67]	.287 [−.712, 1.29]
Interaction	—	—	—	**.141** [.073, .209]	.209 [−.134, .552]	−.172 [−.460, .115]
Linguistic comprehension curvilinear	—	—	—	**.138** [.074, .202]	.050 [−.151, .252]	.014 [−.210, .237]
Word decoding curvilinear	—	—	—	**−.147** [−.198, .096]	**−.324** [−.590, .058]	−.039 [−.336, .258]
Residuals	**.229** [.121, .336]	**.706** [.393, 1.20]	**.520** [.203, .836]	.058 [−.011, .127]	**.615** [.261, .970]	.341 [−.019, .702]

Note. Standardized coefficients; bold values denote $p < .05$.

by the significant curvilinear effects). There was a stronger relationship between listening comprehension and the initial status of reading comprehension for those with better reading comprehension skills. The relationship between word decoding and the initial status of reading comprehension was weaker for those with better reading comprehension skills. Thus, when decoding skills are good, variations in decoding are not of much importance for reading comprehension, whereas the opposite seems to be the case for listening comprehension. A negative curvilinear effect was also found between word decoding and the early growth of reading comprehension. Figure 4 shows this model after removing nonsignificant interactions and curvilinear effects. For reasons of simplicity, the observed variables of the first-order language constructs are not shown in the figure. Additionally, nonsignificant regressions on listening comprehension and word decoding are not shown (but were estimated).

Figure 5a shows the nature of the moderated relationship between listening comprehension and the initial status of reading comprehension: At 1 *SD* above the mean of word decoding, the strength of the path from listening comprehension increases .17 *SD* units, equaling .85 (.68 + .17). At 1 *SD* below the mean of word decoding, the path decreases to .51 (.68 − .17). Thus, the association between listening comprehension and reading comprehension is stronger for good decoders than for poor decoders.

The nature of the curvilinear relationships are illustrated in Figure 5b (initial status on word decoding) and 5c (initial status on listening comprehension). At 1 *SD* above the mean on word decoding, 1 *SD* change in word decoding is associated with a .35 (.49 − .14) *SD* change in the initial status of reading comprehension. At 1 *SD* above the mean on listening comprehension, the path from listening comprehension to the initial status of reading comprehension becomes .81 (.68 + .13) for the initial status of reading comprehension. Because the interaction between listening comprehension and word decoding was significant, we must also take this into account. As an example, at 1 *SD* above the mean on both listening comprehension and word decoding, the path from listening comprehension to the initial status of reading comprehension becomes .98 (.68 + .13 + .17). At 1 *SD* below the mean on both listening comprehension and word decoding, the path from listening comprehension to the initial status of reading comprehension becomes .38 (.68 − .13 − .17). Figure 5d illustrates the curvilinear effect of word decoding on early growth.

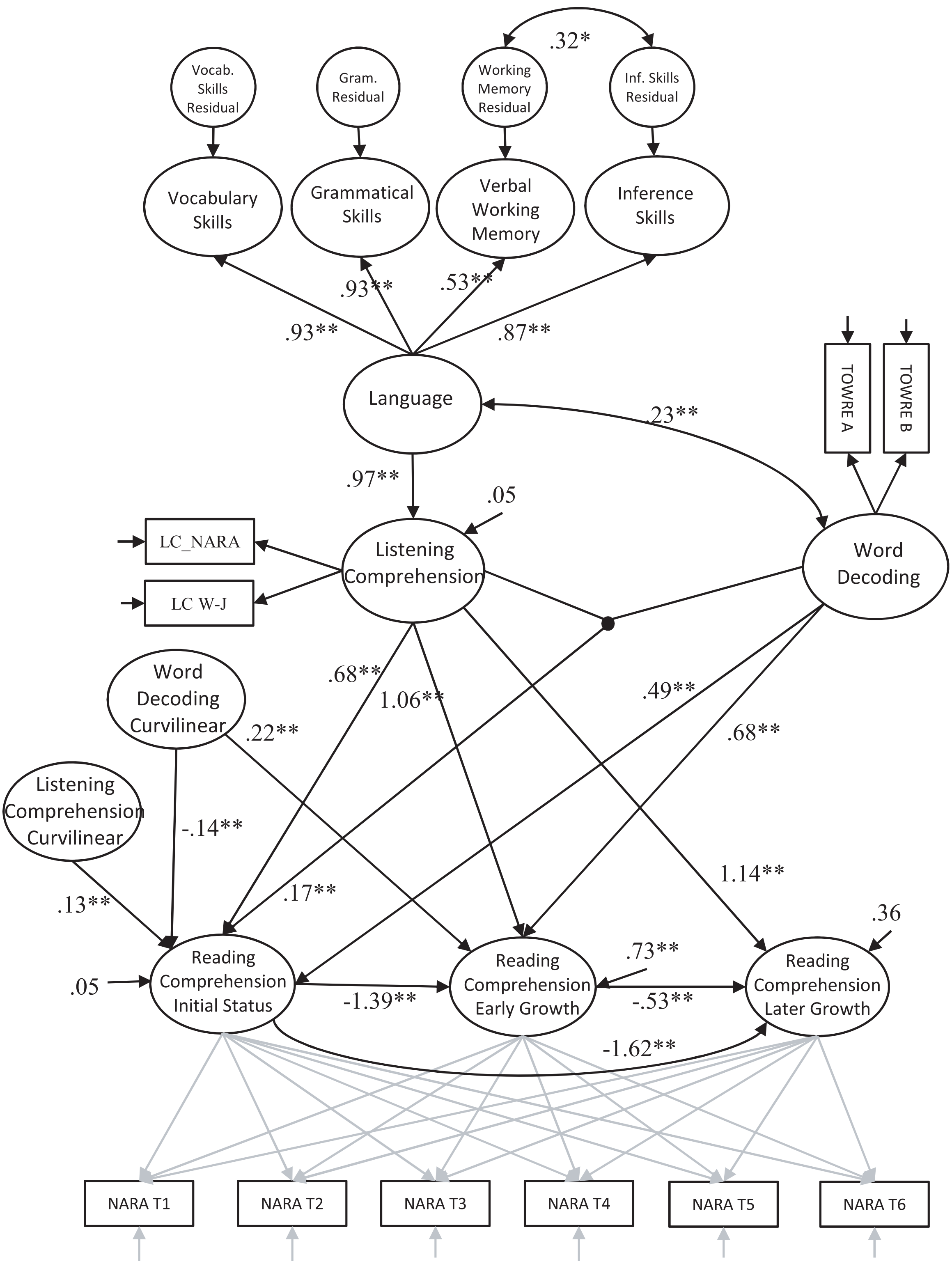

Figure 4. The final structural equation model in which the growth of reading comprehension skills is regressed on listening comprehension and word decoding. Nonsignificant paths from listening comprehension and word decoding are estimated but not shown. The initial status of reading comprehension is also regressed on the interaction between listening comprehension and word decoding plus the curvilinear effects of these two predictors. The early growth of reading comprehension is also regressed on the curvilinear effect of word decoding (other nonsignificant interactions and nonlinear relations are not estimated in this final model). Parameters shown are standardized. For simplification, the observed variables of the first-order language constructs is not shown (see Figure 3). * = $p < .05$; ** = $p < .01$.

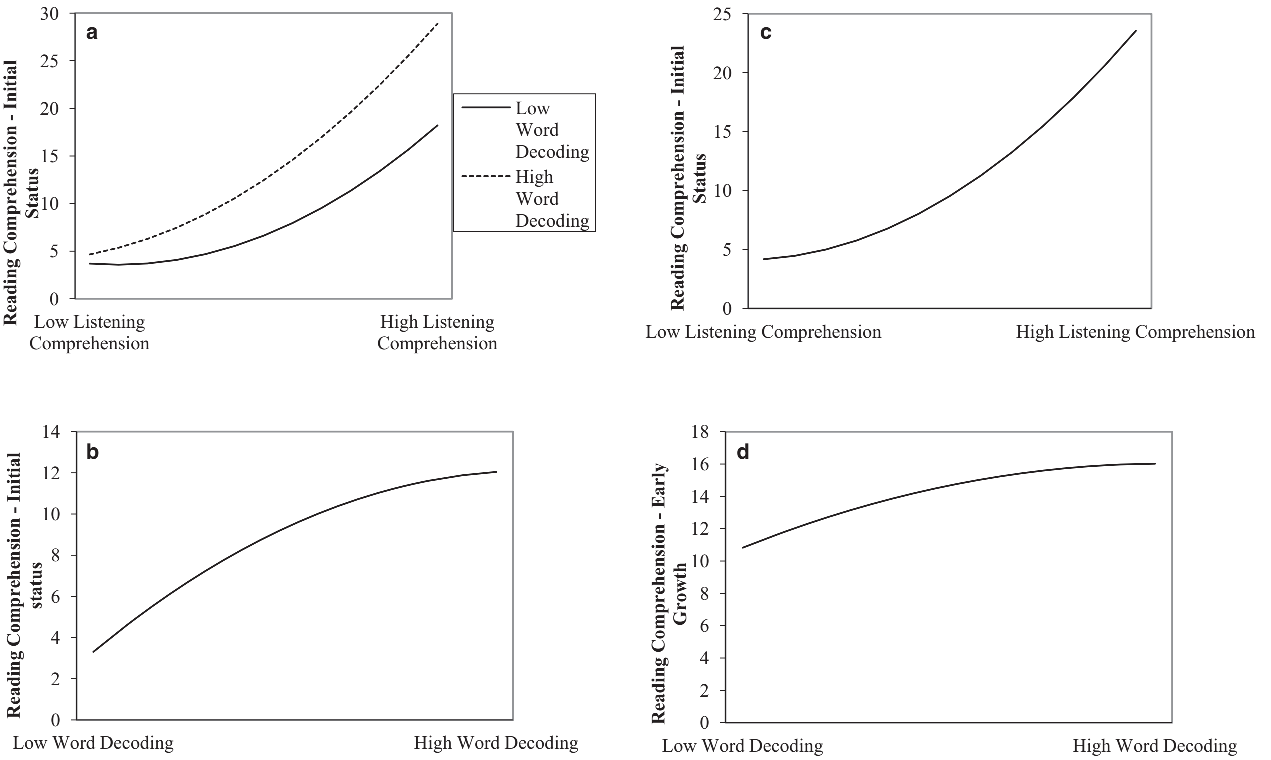

Figure 5. Regression slopes indicating the moderating effect of word decoding on (a) the relationship between listening comprehension and the initial status of reading comprehension, (b) the curvilinear association between word decoding and the initial status of reading comprehension, (c) the curvilinear association between listening comprehension and the initial status of reading comprehension, and (d) the curvilinear association between word decoding and the early growth of reading comprehension.

In this final model with interactions and curvilinear effects included, listening comprehension predicted both the early and the later growth of reading comprehension, whereas word decoding predicted only the early growth. The model explained 95% of the variance in reading comprehension skills in the middle of second grade (initial status), 27% of the early growth and 64% of the later growth.

Simple View or Augmented Simple View?

Finally, we tested whether vocabulary, grammar, verbal working memory, or inference skills predicted the development of reading comprehension skills after the effects of listening comprehension were controlled. We tested this by regressing the three growth factors on the residuals of these four first-order constructs (i.e., on the unique parts of these constructs) one predictor at a time. Table 4 shows the results of these regressions. Only verbal working memory predicted the initial status of reading comprehension skills after controlling for listening comprehension. However, if we use the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995) to correct for the number of possible combinations of predictors and dependent growth constructs ($4 \times 3 = 12$), none of the four constructs are significant predictors of the growth of reading comprehension skills after controlling for the effects of the general language factor. Additionally, in the final model (Figure 4), the language factor (defined by vocabulary, grammar, verbal working memory, and inference skills) explained 95% of the variance in listening comprehension, indicating that this factor is virtually isomorphic with listening comprehension. Regressing the three reading comprehension constructs simultaneously on the common language factor did not improve the model fit, Wald test (3) = 4.174, p = .243, indicating that the influence of the language factor on reading comprehension growth is fully mediated through listening comprehension.

Discussion

This longitudinal study has tested several important hypotheses about the development of reading comprehension across a period of major developmental

Table 4
Growth of Reading Comprehension Regressed on the Unique Parts (Residuals) of the Five Linguistic Comprehension Constructs.

Predictors	Initial status β [95% CI]	Early growth β [95% CI]	Later growth β [95% CI]
Vocabulary skills	.077 [−.051, .205]	−.034 [−.558, .491]	.352 [−.001, .705]
Grammatical skills	.082 [−.128, .293]	.199 [−.128, .526]	.187 [−.219, .593]
Working memory	**.069** [.007, .131]	−.016 [−.223, .191]	.167 [−.092, .425]
Inference skills	.056 [−.073, .186]	−.284 [−.701, .133]	.176 [−.036, .588]

Note. Standardized coefficients; bold value denotes $p < .05$.

change. We found that variations in listening comprehension were almost entirely explained by the variance that is shared by our collection of oral language measures (vocabulary, grammar [syntax and morpheme generation], verbal working memory, and inference skills). The unique parts of these four constructs did not explain additional variance in listening comprehension after controlling for the variance shared by them. The simple view of reading is very strongly supported by our results, because listening comprehension and decoding skills, together with their interaction, explained almost all of the variance in reading comprehension at the beginning of the study. In addition, listening comprehension continued to predict both the early and later growth of reading comprehension, whereas decoding only predicted the early growth for those with poor decoding skills. There was no evidence that vocabulary, grammar, verbal working memory, or inference skills uniquely explained variation in reading comprehension beyond their role as part of this common language factor.

The Structure of Listening Comprehension

We found that variations in listening comprehension were almost entirely explained (95% in the final model) by a factor defined by vocabulary, grammar (syntax and morpheme generation), verbal working memory, and inference skills. This result is at odds with several prior studies using observed variables that found unique contributions to listening comprehension from inference skills and verbal working memory (Kim, 2015, 2016; Lepola et al., 2012). The difference between our study and earlier ones most likely reflects the fact that we have eliminated the problems caused by measurement error by using latent variable models. Theoretically, our finding that verbal working memory correlates so well with other measures of language ability is consistent with the view that these tasks essentially reflect differences in the underlying language skills on which they depend (Allen & Hulme, 2006; MacDonald & Christiansen, 2002) rather than the idea that verbal working memory is a separate "language-learning device" (Baddeley, Gathercole, & Papagno, 1998; Gathercole, Willis, Emslie, & Baddeley, 1992).

Additive, Product, and Nonlinear Effects of Listening Comprehension and Word Decoding Explain the Growth of Reading Comprehension

As predicted by the simple view of reading, the relationship between listening comprehension and reading comprehension in Grade 2 varied as a function of decoding skills. For good decoders, variations in listening comprehension are more predictive of how well they understand written texts than for poor decoders. Without adequate levels of decoding, oral language comprehension skills cannot be engaged to allow the comprehension of a written text.

The curvilinear relationship between decoding and reading comprehension shows that individual variations in word decoding skills are less important among good decoders than poor decoders. Once decoding skills are sufficient to decode a text, further improvements in decoding skill are irrelevant. However, when decoding skills are poor, slightly better decoding skills can lead to better understanding of a text. A similar relationship was found between word decoding and the early growth of reading comprehension skills. For good decoders, individual variations in decoding skills have little effect on the development of reading comprehension skills. In contrast, for poor decoders, individual variations in decoding skills may play an important role in how fast their reading comprehension skills develop. These findings are a further illustration of the bottleneck function that word decoding seems to have when a child is trying to comprehend a written text. In contrast, listening comprehension continued to predict growth in both early and later reading comprehension skills for all participants and no interaction was found here. These results underline the long-lasting and important role that listening comprehension plays in the development of reading comprehension skills.

Augmented Simple View

According to the augmented simple view of reading, certain component language skills (e.g., inference making and verbal working memory skills) make a direct contribution to reading comprehension in addition to any role they might play in fostering listening comprehension. In contrast, the simple view of reading posits that these component language skills only affect reading comprehension via their effects on listening comprehension. Contrary to some earlier studies using observed variables (e.g., Geva & Farnia, 2012; Oakhill & Cain, 2012), we found no support for the augmented simple view because the effects of component language skills on reading comprehension were entirely accounted for by their effects on listening comprehension. Similar findings were obtained by Kim (2015), in a concurrent study. We believe measurement error is a likely reason for the discrepancy between our study and many of the prior studies that found differentiated effects of vocabulary, grammar, verbal working memory, and inference skills on reading comprehension skills (see Cole & Preacher, 2014).

Strengths and Limitations

There are several strengths of the current study. We have a relatively large sample who were assessed repeatedly over a period of significant developmental change (second–seventh grades) using a rich test battery. In addition, the use of latent variables for all constructs allowed us to control for measurement error. Controlling for measurement error is particularly important when assessing the relative importance of different predictors and when testing for interactions between predictors.

The majority of earlier studies that have related the development of reading comprehension to a large set of predictors over time have been conducted with children learning to read in an opaque orthography—English (Hjetland, Brinchmann, Scherer, & Melby-Lervåg, submitted). The current study involved children learning to read in the relatively consistent Norwegian orthography, where decoding is mastered quickly (see also Caravolas, Lervåg, Defior, Malková, & Hulme, 2013). Because of this, variations in language comprehension start to play a role in limiting reading comprehension relatively earlier than for children learning to read in English (in the current sample around second grade, in English speaking samples third–fourth grades, see Storch & Whitehurst, 2002). Children learning to read in consistent orthographies seem to surpass English children's decoding skills quickly after the start of formal reading instruction even though formal reading instruction in English starts 1–2 years earlier (see Caravolas et al., 2013 for a comparison between Spanish and Czech vs. English readers). However, despite the difference between orthographies, the patterns of development found here correspond closely with earlier findings from English-speaking samples (Hjetland et al., submitted). Therefore, the pattern of predictors of reading comprehension seems to be relatively consistent across orthographies that differ in orthographic consistency. The current model explains an impressive 95% of the variation in listening comprehension, and 96% of the variance in the initial level of reading comprehension, which suggests that there is little variance left to explain in these constructs. It is still possible however that other constructs not included in the current study might play a role in the development of reading comprehension. For example, Kim (2016) found that attention explained variance in listening comprehension, and this was not assessed here. Background knowledge has also been suggested to be important for reading comprehension (Cromley & Azevedo, 2007; Kintsch, 1988; McNamara & Kintsch, 1996; Perfetti & Stafura, 2014). However, the effects of background knowledge on reading comprehension may depend on differences in vocabulary knowledge, because recent studies with children that have suggested that these are not separable constructs (see Kim, 2016). The percentage of variance explained in the early and later growth of reading comprehension was less than in the initial level of reading comprehension suggesting that unmeasured construct might play a more important role here. However, the residual variance in the later growth of reading comprehension was not significantly different from zero indicating that little if any variance in growth remains to be explained.

Notwithstanding the very high proportion of variance in both reading and listening comprehension accounted for by the models presented here, there is always the possibility that if other variables had been used to represent the constructs (e.g., working memory and inference skills), the results might have been different. Because we used a verbal working memory measure that is closely related to reading comprehension, it does not seem plausible that a nonverbal working memory task would have produced better predictions of reading comprehension. Similarly, the inference task we used focused on word learning from context, because we wanted to see whether such language inference skills were unique predictors of both listening comprehension and reading comprehension beyond the general language skills that such tasks contain. It might be, however, that other tasks more directly associated with making inferences from written text (e.g., making inferences within sentences and across parts of texts) would have been even more successful in predicting reading comprehension skills. In addition, different measures of reading comprehension vary in how dependent they are on decoding and linguistic comprehension, respectively (e.g., Keenan et al., 2008). Therefore, the nature of the reading comprehension test can affect the results. The NARA used here seems to tap both listening comprehension and decoding, but decoding less than comprehension.

Summary and Conclusions

In summary, our findings indicate that multiple language-related skills are involved in listening comprehension, which in turn is a powerful influence on the development of reading comprehension. Our results give very strong support to the simple view of reading and clarify a number of theoretical issues concerning the relationship between oral language and reading comprehension skills. Our findings also have implications for how to prevent and ameliorate reading comprehension problems. First, for children with poor decoding skills, it seems decoding can be a bottleneck for the development of reading comprehension—interventions to improve decoding in those with poor skills can therefore be expected to lead directly to improvements in reading comprehension. Even fairly minimal improvements in decoding for poor decoders may have functionally important implications for reading comprehension for this group. At a more general level, our results suggest that interventions that also focus on a broad set of oral language skills, including grammar, syntax, narrative skills, and inference making are most likely to be effective in helping children to develop adequate reading comprehension skills. There are now a handful of randomized controlled trials that have examined the effects of interventions to improve language comprehension (e.g., Clarke et al., 2010; Fricke, Bowyer-Crane, Haley, Hulme, & Snowling, 2013; Rogde, Melby-Lervåg, & Lervåg, 2016; see also Melby-Lervåg & Lervåg, 2014). Several studies support the claim that interventions can improve language comprehension skills in young children (Fricke et al., 2013; Rogde et al., 2016) and improvements in oral language skills appear to lead directly to improvements in reading comprehension both in younger (Fricke et al., 2013) and older children Clarke et al., 2010). Such findings provide strong support for the simple view of reading, and for the causal theory that the development of reading comprehension is dependent on underlying oral language skills.

References

Aarnoutse, C., van Leeuwe, J., & Verhoeven, L. (2005). Early literacy from a longitudinal perspective. *Educational Research and Evaluation*, *11*, 253–275. https://doi.org/10.1080/08993400500101054

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Verbal Learning & Verbal Behavior*, *55*, 64–88. https://doi.org/10.1016/j.jml.2006.02.002

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173. https://doi.org/10.1037/0033-295X.105.1.158

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, *57*, 289–300. doi:10.2307/2346101

Bishop, D. V. (2009). *Test for Reception of Grammar* (TROG-2, Norwegian version). London, England: Pearson Education.

Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology*, *75*, 189–201. https://doi.org/10.1348/000709904X22674

Cain, K., & Oakhill, J. (2006). Assessment matters: Issues in the measurement of reading comprehension. *British Journal of Educational Psychology*, *76*, 697–708. https://doi.org/10.1348/000709905X69807

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Cain, K., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, *30*, 681–694. https://doi.org/10.1017/S0305000903005713

Caravolas, M., Lervåg, A., Defior, S., Malková, G. S., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science*, *24*, 1398–1407. https://doi.org/10.1177/0956797612473122

Chen, R. S., & Vellutino, F. (1997). Prediction of reading ability: A cross-validation study of the simple view of reading. *Journal of Literacy Research*, *29*(1), 1–24. https://doi.org/10.1080/10862969709547947

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., . . . Olson, R. K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. *Journal of Experimental Psychology General*, *141*, 470–488. https://doi.org/10.1037/a0027375

Clarke, P. J., Snowling, M. J., Trulove, E., & Hulme, C. (2010). Ameliorating children's reading comprehension difficulties: A randomized controlled trial. *Psychological Science*, *21*, 1106–1116. https://doi.org/10.1177/0956797610375449

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*, 300–315. https://doi.org/10.1037/a0033805

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*, 311–325. https://doi.org/10.1037/0022-0663.99.2.311

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*, 277–299. https://doi.org/10.1207/s1532799xssr1003_5

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading*, *6*(1), 51–77. https://doi.org/10.1207/S1532799XSSR0601_03

Dufva, M., Niemi, P., & Voeten, M. J. (2001). The role of phonological memory, word recognition, and comprehension skills in reading development: From preschool to grade 2. *Reading and Writing*, *14*, 91–117. https://doi.org/10.1023/A:1008186801932

Dunn, L. M., & Dunn, L. M. (1997). *PPVT–III: Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.

Florit, E., Roch, M., & Levorato, M. C. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes*, *48*, 119–138. https://doi.org/10.1080/0163853X.2010.494244

Foorman, B. R., Koon, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). Examining general and specific factors in the dimensionality of oral language and reading in 4th–10th grades. *Journal of Educational Psychology*, *107*, 884–899. https://doi.org/10.1037/edu0000026

Fricke, S., Bowyer-Crane, C., Haley, A., Hulme, C., & Snowling, M. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry*, *54*, 280–290. doi:10.1111/jcpp.12010

Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological Methods*, *2*, 235–247. https://doi.org/10.1037/1082-989X.2.3.235

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*, 74–111. https://doi.org/10.3102/0034654313499616

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, *28*, 887–898. https://doi.org/10.1037/0012-1649.28.5.887

Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL. *Reading and Writing*, *25*, 1819–1845. https://doi.org/10.1007/s11145-011-9333-8

Gjessing, H., & Nygaard, H. (1995). *Illinois Test of Psycholinguistic Abilities (ITPA). Håndbok, Norsk utgave, femte opplag*. (Manual, Norwegian version, fifth edition) Oslo, Norway: Skolepsykologi—materiellservice.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7* (1), 6–10. https://doi.org/10.1177/074193258600700104

Hjetland, H. N., Brinchmann, E. I., Scherer, R., & Melby-Lervåg, M. (submitted). Preschool predictors of later reading comprehension ability: A systematic review. *Campbell Systematic Reviews*.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*, 127–160. https://doi.org/10.1007/BF00401799

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

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, *21*, 85–97. https://doi.org/10.1080/02702710050084428

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, *12*, 281–300. https://doi.org/10.1080/10888430802132279

Kendeou, P., Van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, *101*, 765–778. https://doi.org/10.1037/a0015956

Kim, Y. S. (2015). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development*, *86*, 128–144. https://doi.org/10.1111/cdev.12293

Kim, Y. S. (2016). Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology*, *141*, 101–120. https://doi.org/10.1016/j.jecp.2015.08.003

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kirby, J. R., & Savage, R. S. (2008). Can the simple view deal with the complexities of reading? *Literacy*, *42*, 75–82. https://doi.org/10.1111/j.1741-4369.2008.00487.x

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, *18*, 146–154. https://doi.org/10.1111/desc.12202

Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4–6-year-old children. *Reading Research Quarterly*, *47*, 259–282. https://doi.org/10.1002/rrq.020

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*, 612–620. https://doi.org/10.1111/j.1469-7610.2009.02185.x

Limbird, C. K., Maluch, J. T., Rjosk, C., Stanat, P., & Merkens, H. (2014). Differential growth patterns in emerging reading skills of Turkish-German bilingual and German monolingual primary school students. *Reading and Writing*, *27*, 945–968. https://doi.org/10.1007/s11145-013-9477-9

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54. https://doi.org/10.1037/0033-295X.109.1.35

McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*, 247–288.

Melby-Lervåg, M., & Lervåg, A. (2014). Effects of educational interventions targeting reading comprehension and underlying components. *Child Development Perspectives*, *8*, 96–100. https://doi.org/10.1111/cdep.12068

Melby-Lervåg, M., Lervåg, A., Lyster, S.-H., Klem, M., Hagtvet, B., & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science*, *23*, 1092–1098. https://doi.org/10.1177/0956797612443833

Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

Näslund, J. C. (1990). The interrelationships among preschool predictors of reading acquisition for German children. *Reading and Writing*, *2*, 327–360. https://doi.org/10.1007/BF00419387

Neale, M. D. (1997). *Neale analysis of reading ability: Second revised* (British ed.). London, UK: NFER-Nelson.

NICHD Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology*, *41*, 428–442. https://doi.org/10.1037/0012-1649.41.2.428

Oakhill, J., & Cain, K. (2000). Children's difficulties in text comprehension: Assessing causal issues. *Journal of Deaf Studies and Deaf Education*, *5*, 51–59. doi:10.1093/deafed/5.1.51

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*, 91–121. https://doi.org/10.1080/10888438.2010.529219

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Pickering, S., & Gathercole, S. (2001). *Working Memory Test Battery for Children (WMTB-C) manual*. London, UK: The Psychological Corporation.

Protopapas, A., Mouzaki, A., Sideridis, G. D., Kotsolakou, A., & Simos, P. G. (2013). The role of vocabulary in the context of the simple view of reading. *Reading & Writing Quarterly*, *29*, 168–202. https://doi.org/10.1080/10573569.2013.758569

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development*, *86*, 159–175. https://doi.org/10.1111/cdev.12292

Rogde, K., Melby-Lervåg, M., & Lervåg, A. (2016). Improving the general language skills of second-language learners in kindergarten: A randomized controlled trial. *Journal of Research on Educational Effectiveness*, *9*(sup1), 150–170, https://doi.org/10.1080/19345747.2016.1171935

Roth, F. P., Speece, D. L., Cooper, D. H., & De la Paz, S. (1996). Unresolved mysteries: How do metalinguistic and narrative skills connect with early reading? *Journal of Special Education*, *30*, 257–277. https://doi.org/10.1177/002246699603000303

Seigneuric, A., & Ehrlich, M. (2005). Contribution of working memory capacity to children's reading comprehension: A longitudinal investigation. *Reading and Writing*, *18*, 617–656. https://doi.org/10.1007/s11145-005-2038-0

Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 173–182). New York, NY: Guilford.

Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology*, *107*, 321–331. https://doi.org/10.1037/a0037769

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, *38*, 934–947. https://doi.org/10.1037/0012-1649.38.6.934

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency (TOWRE)*. Austin, TX: Pro-Ed.

Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing*, *25*, 1805–1818. https://doi.org/10.1007/s11145-011-9346-3

Wechsler, D. (2003). *Wechsler intelligence test for children*. San Antonio, TX: The Psychological Corporation.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson tests of achievement*. Itasca, IL: Riverside.

Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. New York, NY: Cambridge University Press.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Estimated Correlations Between All Variables at All Time Points

Data S1. Examples of the Inference Tasks (Translated From Norwegian)