2017

Modeling the early language trajectory of language development and its relation to poor reading comprehension

Yaacov Petscher, Laura Justice, Tiffany Hogan and Andrew Mashburn

This is the peer reviewed version of the following article: Modeling the Early Language Trajectory of Language Development When the Measures Change and Its Relation to Poor Reading Comprehension, which has been published in final form at http://dx.doi.org/10.1111/cdev.12880.

Modeling the Early Language Trajectory of Language Development and Its Relation to Poor Reading Comprehension

Petscher, Y., Justice, L., Hogan, T., & Mashburn, A. (in press). Modeling the early language trajectory of language development and its relation to poor reading comprehension. *Child Development.*

Abstract

This study examined change in early language comprehension from 15 to 54 months for fifth grade typical readers (*n* = 35), poor decoders (*n* = 11), or poor comprehenders (*n* = 16) who were ascertained at birth in 1991 in a nationally representative study of early child-care experiences. Measures of language comprehension were captured across early childhood for the 72 children, and these measures changed at each measurement time-point; therefore, data were transformed to ranks. Multiple group quasi-simplex and latent growth models were used to examine children's relative rank change. Results showed that future poor comprehenders significantly declined in language comprehension over time relative to future poor decoders and typical readers, who gradually improved. Findings suggest that deficits in early language contribute to reading difficulties. Efforts to improve language skills as a means to improve reading comprehension, particularly for poor comprehenders, hinge upon the perspective that language weaknesses are a causal contributor to reading difficulties.

Keywords: Language development, Matthew effects, multiple group SEM

Modeling the Early Trajectory of Language Development and Its Relation to Poor Reading Comprehension

Children's acquisition of language over the early years of life represents one of the most well-documented developmental phenomenon within the scientific literature. This literature includes substantial efforts to identify both universals in language acquisition across the many world languages (e.g., Givon, 2013) as well as individual differences that serve to differentiate among children in their acquisition of specific linguistic forms and functions (Morgan-Short et al., 2014). The latter body of work is particularly influential to identifying children who do not acquire language skills at the same rate or in the same manner as would be expected given their age and their cultural background. For instance, children acquiring English who do not combine words to form two-word utterances by two years of age are often viewed as 'late talkers' (Ellis et al., 2015), a subset of whom will not outgrow these early lags and will experience pervasive, ongoing difficulties with language skill, characteristic of a developmental language impairment. Children with impaired language skills lag behind their same-age peers in achieving important milestones of language acquisition, such as producing complex sentence structures and inflecting verbs to mark tense (Schuele & Dykes, 2005). Such limitations affect spoken discourse and, as is increasingly documented, their ability to process and comprehend written text as well (Tambyraja et al., 2015). Research demonstrating the reading-comprehension difficulties of children with developmental language impairment is important to improving our understanding of the role of language skill in reading for meaning.

To this end, theoretical models of reading development (Hoover & Gough, 1990) as well as empirical tests of these models (Catts, Adlof, & Ellis Weismer, 2006) provide evidence for the prospective longitudinal relations of children's pre-reading language skills to their ability to read for meaning in the later grades (Catts, Adlof, & Ellis Weismer, 2006; Catts, Herrera, Nielsen, & Bridges, 2015; Duff, Tomblin, & Catts, 2015; McKean et al., 2015; Murphy, Language and Reading Research Consortium, Farquharson, 2015). For instance, research findings have shown that children's language skills as measured at three years of age directly contribute to individual differences in reading comprehension at third grade [National Institute of Child Health and Human Development Early Child Care Research Network (NICHD ECCRN), 2005]. The NICHD ECCRN study showed children's early language skills contributed both directly and indirectly (the latter through their impact on phonological awareness and alphabet knowledge) to third-grade reading comprehension. Longitudinal research (see also National Early Literacy Panel, 2008) clarifies that children's early language skills are not only important precursors to skilled reading comprehension, but that their contribution to skilled reading should be differentiated from those skills that serve as precursors to decoding, such as phonological awareness (Catts, Adlof, Hogan, & Ellis Weismer, 2005; Hoover & Gough, 1990; Storch & Whitehurst, 2002).

The "simple view of reading" (SVR) provides a useful framework for conceptualizing the role of language skill in skilled reading (Gough & Tunmer, 1986). The SVR specifies that reading comprehension (R) is the product of decoding (D) and listening comprehension (C), with listening comprehension considered to be synonymous with language comprehension (Language and Reading Research Consortium, in press). Put simply, the decoding or word recognition (D) component of this model translates print into language and the comprehension (C) component makes sense of this linguistic information. Numerous studies show that word recognition and language comprehension are relatively independent of each other, but both contribute significantly and uniquely to reading comprehension (e.g., Aaron, Joshi, & Williams, 1999; de

Jong & van der Leij, 2002; Hoover & Gough, 1990). Further, studies show that the contribution of word recognition and language comprehension to reading comprehension varies across grades (Aaron et al., 1999; Catts et al., 2005). In the early grades, reading comprehension is heavily dependent on emerging word reading skills; however, as these become automatized, language skills serve as a more critical determinant of one's ability to read for meaning. To illustrate, a recent cross-sectional study involving 371 first, second, and third grades examined the contribution of word recognition and language comprehension at each grade to a latent variable reflecting reading comprehension (Language and Reading Research Consortium, 2015). Whereas word recognition and language comprehension explained 64% and 24% of the variance, respectively, in reading comprehension at first grade, at third grade the contributions of word recognition and language comprehension were 21% and 61%, respectively. These findings show the prominent, unique contribution of oral language skill to reading comprehension in the later primary grades.

Improved understanding of the role that language skills play in skilled reading comprehension has contributed substantially to knowledge about the reading difficulties that characterize a unique group of poor readers referred to as *poor comprehenders* (Catts, Hogan, & Adlof, 2005; Nation, 2005; Nation, Clarke, & Snowling, 2002). Referred to as comprehension-specific reading disability, poor comprehenders represent an estimated 15% of poor readers. This subgroup of poor readers exhibit accurate and fluent word-reading skills and show normal intelligence, yet experience extreme difficulty comprehending what they read (Stothard & Hulme, 1995; Yuill & Oakhill, 1991). Theoretical and empirical evidence showing the direct contributions of language skills to the ability to read for meaning suggests that poor comprehenders' reading difficulties likely reflect an underlying developmental language impairment (Hulme & Snowling, 2009).

A developmental language impairment, referenced previously, is a disability in which a child exhibits significant lags in language development in the absence of any obvious causal factors, such as hearing loss, intellectual disability, or autism; thus, a key feature of developmental language impairment is that the cause of the language impairment is of unknown origin (Bishop, 2014). Other terms used to describe this condition include primary language impairment and specific language impairment, although these terms often describe children whose language difficulties occur in absence of significant intellectual disability (Bishop, 2014). Children who are poor comprehenders in the later primary grades exhibit difficulties with language skill that are similar to children with developmental language impairment. For instance, they perform more poorly than typical children on measures representing complex syntax (e.g., processing of anaphoric references in discourse-level tasks; Megherbi & Ehrlich, 2005) and higher-level semantic processing (e.g., recognizing relations among synonyms; Nation & Snowing, 1998). They also perform poorly on measures of verbal working memory, showing difficulties inhibiting irrelevant information (Cain, 2006).

Poor performance on such linguistic tasks are typically interpreted as showing that poor comprehenders' deficits in reading comprehension reflect deficits in the subsystems of language, including grammar and semantics, leading some to argue that poor language skills are *causal* to poor reading comprehension (Hulme & Snowling, 2009). Further, evidence suggests that poor language skills may be a heritable component of reading difficulties. Results of a recent meta-analysis showed that infants and toddlers with immediate family members with reading disabilities had significantly poorer vocabulary and grammar skills than children without such familial risk (Snowling & Melby-Lervåg, 2016).

Do Poor Language Skills "Cause" Poor Reading Comprehension?

To date, there is little experimental evidence available regarding the potentially deterministic role of poor language skills to reading comprehension problems. Thus, we cannot determine that presence of a developmental language impairment is a causal contributor to the reading difficulties of poor comprehenders. However, there is a large volume of descriptive work identifying the specific linguistic deficits that appear to contribute most prominently to the poor reading comprehension observed among poor comprehenders, such as the ability to make inferences and to monitor one's comprehension when reading (Cain, Oakhill, & Bryant, 2004; Oakhill, 1982, 1983). The consistently observed deficits in the performance of poor comprehenders on linguistic tasks, as compared to poor decoders and typical readers, has contributed to a general perspective situating poor language skills as deterministic to poor reading comprehensions. Nonetheless, such studies are generally cross-sectional and largely involve examining concurrent relations between linguistic skills (e.g., ability to make inferences) and children's ability to read for meaning. As Hulme and Snowling (2009) point out, longitudinal research is "badly needed to investigate the causes as well as the consequences of reading comprehension impairments" (p. 120). Longitudinal research that models language development over multiple time points would be informative to understanding the pathway by which impairments in spoken language skill lead to impairments in reading comprehension.

To our knowledge, the peer-review literature contains only three longitudinal studies seeking to determine the extent to which language deficits may indeed play a causal role in comprehension-specific reading disability. While there are numerous longitudinal studies of language development and its relation to future reading skill, these particular studies are focused longitudinally on understanding the historical language performance of children who exhibit poor reading comprehension. Thus, these studies involved *retrospective* examination of the language skills of poor comprehenders relative to other reading subgroups (Catts et al., 2006; Justice, Mashburn, & Petscher, 2013; Nation, Cocksey, Taylor, & Bishop, 2010). Collectively, these studies provide suggestive evidence that the language weaknesses of poor comprehenders are apparent *prior* to reading instruction and thus may be viewed as causal antecedents of the reading problems that emerge among poor comprehenders. Catts et al (2006) examined the language-comprehension skills of 57 eighth-grade poor comprehenders retrospectively to age five years, drawing comparisons against future poor decoders (n = 27) and future typical reader (n = 98). At five years, children who would become poor comprehenders had significantly lower skills in language comprehension compared to both comparison groups on a composite language measure; the magnitude of these effects were large, particularly in relation to future typical readers (d = 1.37) but also for future poor decoders (d = 0.76).

Important to note within the Catts et al (2006) report is that on two of three language measures administered, representing vocabulary and discourse skills, the language skills of future poor comprehenders appeared to substantially *decline* over time (from kindergarten to fourth grade), both in absolute value and as compared to the other reader subgroups (i.e., typical readers and poor decoders). As this study did not actually examine growth over time, but rather used analysis of variance to compare language scores at three individual time points across the three comparison groups, it is difficult to draw conclusions regarding the trajectory of language growth among the poor comprehenders, particularly the possibility of deteriorating language abilities with time. Nonetheless, we can theorize that the language skills of children who are poor comprehenders might weaken over time in light of decreasing exposure to complex texts (with which they would struggle to comprehend particularly in relation to peers), which are a critical source of language stimulation across the primary grades (see Adams, 2010). Indeed, Hulme and Snowling (2009) have speculated that children with poor language skills may become poorer over time in their reading ability relative to other children due to constraints imposed by their language impairment. For instance, when reading, these children are not skilled at inferring the meanings of words from context, which results in a comparative disadvantage in learning new words from the written context relative to other children. Thus, when other children are reading often and fueling their vocabularies from the written context, children with poor language skills are falling further behind, consistent with a Matthew Effect (the rich get richer while the poor get poorer; Cain & Oakhill, 2011; Morgan, Farkas, & Wu, 2011). Thus, while poor language skills are considered a critical antecedent of the reading problems of poor comprehenders, they may also be considered a key contributor to the ongoing lags in reading growth observed among these children in relation to others.

In addition to the longitudinal research of Catts and colleagues (2006), Nation and colleagues (2010) similarly used longitudinal methods to examine the language skills of 15 eight-year-old poor comprehenders retrospectively at age five relative to 15 eight-year-old typical readers. In their study, the two reader groups were matched for word-reading skill and thus differed only with respect to reading comprehension. Upon retrospective examination of language skills at age five, the future poor comprehenders were shown to have significantly poorer language skills than the typical readers on measures of both grammar and vocabulary; similar to Catts et al.'s findings, effects were large to very large in magnitude (d =.64 for vocabulary and 1.05 for grammar and vocabulary). Although this study did not examine language trajectories per se, there was no evidence that the language skills of future poor comprehenders declined over time in relation to typical readers, as observed by Catts and his colleagues and discussed by Hulme and Snowling (2009).

Finally, in the only study of which we are aware that involved retrospective assessment of poor comprehenders' language skills prior to age five years, Justice, Mashburn, and Petscher (2013) examined the language skills of 16 fifth-grade poor comprehenders at 15, 24, 36, and 54 months as compared to 11 fifth-grade poor decoders and 35 typical readers. Comparisons of the three groups at each of these time points showed that future poor comprehenders exhibited lower levels of language skills at every assessment point from 24 to 54 months and across every measure as compared to both future poor decoders and future typical readers. Of additional note, the poor comprehenders' early language skills appeared to diverge over time in relation to poor decoders and typical readers, with the greatest divergence occurring at 54-months in language comprehension. The study findings suggested that the language skills of future poor comprehenders changed over time: the median percentile ranking of poor comprehenders on measures of language comprehension changed from the 23rd percentile at 15 months to the 8th percentile at 54 months. Though these ranks are based on different measures and norming samples, the general indication is that early language skills for future poor comprehenders did not remain static. Similar to Catts et al.'s (2005) work examining the language skills of poor comprehenders from kindergarten through fourth grade, the findings of Justice and colleagues (2013) suggested that both in absolute terms and in comparison to other reader groups, children with comprehension-specific reading problems may have not only a substantial weakness in language ability, but that these skills declined over time in relation to other children.

The Matthew Effect: A Salient Phenomenon in Language Trajectories of Poor Comprehenders?

The Matthew Effect is an oft-described developmental phenomenon in which children who have weak skills in a certain area (e.g., math, reading) increasingly fall behind children with strong skills in a certain area. Dependent on the statistical model used, the Matthew Effect is typically characterized by increasing variance over time (Zumbo, 1999) as well as a positive correlation between children's growth over time and their initial status or skill levels, thus showing that children with higher initial skills gain more over time compared to those with lower initial skills (Bast & Reistma, 1997). With respect to the role of oral language deficits to the reading problems of poor comprehenders, findings reported in Catts et al (2006) and Justice et al (2013) both implied that a Matthew Effect may be operating when examining growth in language comprehension over time for children who become poor comprehenders as compared to other children. That is, the results of these two studies suggested that the language skills of poor comprehenders lagged increasingly behind those of other children from the toddler and preschool years (Justice et al., 2013) and the early primary grades (Catts et al., 2006).

Even in the later grades, children with comprehension-specific problems appear to continue to fall behind other children: recent work by Cain and Oakhill (2011) provide further evidence of a Matthew Effect specific to vocabulary, encompassing both reading vocabulary (sight words) and spoken vocabulary. Poor comprehenders as observed from the early primary grades (7-8 years of age) to adolescence (15-16 years of age), grew less in vocabulary skills over time as compared to typical readers. It may be that this phenomenon is a bit of the culprit in the fate of the poor comprehender, such that many of these children may go unnoticed at school (Hulme & Snowling, 2009), because (a) their language skills are poorer than typical children but not so poor as to result in a diagnosis of developmental language disorder (Nation, Clarke, Marshall, & Durand, 2004) and (b) their language problems are best represented dynamically, based on comparison of growth over time to other children.

Such circumstances suggest the need to investigate more thoroughly the early language trajectories of later-diagnosed poor comprehenders, particularly with respect to how early trajectories may appear in relation to other children who go on to show reading difficulties that do not affect comprehension (poor decoders) and who are typical readers. While it is recognized that children with developmental language problems exhibit significant lags in the attainment of language skills as compared to typically developing children (indeed, this is the governing "phenotype" of this disorder), a potentially marked characteristic of this disorder is that these children may exhibit a negative divergence from typical children in their language growth over the course of development. Although there have been mixed findings regarding the nature of growth in language skills among children with developmental language disorders, a recent and rigorous assessment of growth trajectories for children with developmental language disorders from kindergarten to fifth grade provided some support for a Matthew Effect. Morgan and his colleagues (2011) specifically examined the possibility of Matthew Effects for children with language impairment, children with learning disabilities (LD), and typical readers over four available time-points using growth-curve modeling when controlling for sociodemographic variables such as age, race, gender, and socioeconomic status. Growth models consistently showed that children with language impairment experienced less reading growth over time relative to typical readers, thus falling further behind over time in reading skill. This Matthew Effect phenomena observed in reading development over time for children with language impairment during the primary grades converges with observations of a Matthew Effect among poor comprehenders during middle and high school (Cain & Oakhill, 2011).

It is possible that the Matthew Effect is a salient phenomenon that differentiates the early language trajectories of young children who will go on to become poor comprehenders as compared to children who will become typical readers or even those with other typologies of reading disabilities (e.g., poor decoding skills/dyslexia). That is, a defining characteristic of children with comprehension-specific reading problems may be not only that they perform less well than other children on linguistic tasks at a given point in time (e.g., Oakhill, 1982, 1983) but that their language trajectory lags increasingly behind their peers over time (Cain & Oakhill, 2011; Morgan et al., 2011). Likely, this occurs due to the neurobiological underpinnings of the disorder, such that children with developmental language impairment acquire language skills more slowly than other children (Windsor & Hwang, 1999), but also due to environmental factors, such as engaging less consistently during linguistic-based interactions with adults (Skibbe, Moody, Justice, & McGinty, 2008).

Aims and Specific Contributions of the Present Study

The primary aim of the present study was to examine the early language trajectories of children who would go on to become poor comprehenders at fifth grade as compared to children who would become typical readers and poor decoders. Specifically, our interest was to determine whether the Matthew Effect suggested in prior published work (Catts et al., 2006; Justice et al., 2013) served to characterize the early language trajectories of poor comprehenders. Neither of these prior studies examined *growth trajectories* over time, but instead conducted comparisons of language skills at specific points in time for poor comprehenders and other groups of children because the measures of language skills in these studies changed across time-points.

Consequently, it is not possible to discuss definitively whether the language trajectories of poor comprehenders differed in any meaningful way from those of other children.

To address our primary aim, we used a two-pronged approach. The first approach evaluated the effects of one measurement occasion on the next, whereas the second approach studied the extent to which children's relative rank ordering in language comprehension changed over time. For the first approach, we used quasi-simplex models, which represent a conventional methodology to assess longitudinal dependencies in a series of data where the measures change. However, they are limited in that only the first-order autoregressive effects may be examined. Such approaches are ill-equipped to assess average change over multiple waves. Thus, in the present study we also applied an innovative statistical approach that allows one to measure change over time in the context of changing measures. An important feature of this approach is that it does not require large sample sizes (Lloyd et al., 2009), which are difficult to obtain and retain in longitudinal samples of children with specific deficits that have small incidence rates. This statistical approach holds promise for increasing our understanding of the language trajectories of poor comprehenders, as a major factor inhibiting longitudinal study of these youngsters pertains to variations in how children's language skills are measured over time. It is not uncommon for longitudinal studies of language development to involve use of different measures across waves, given that observable characteristics of children's language skills appear to change fundamentally with development; for instance, children who are under two years of age typically do not use complex syntax (and thus one would not assess this aspect of language), whereas children over three years of age commonly do (Vasilyeva, Waterfall, & Huttenlocher, 2008). The changing nature of language development thus influences the types of measures used over time, and as such it is common practice for longitudinal studies to adopt different measures of language across different waves of data collection. The changing of measures over time to assess children's language skills precludes the use of traditional modeling of growth trajectories. Growth modeling has typically been applied to longitudinal data when the measures remain completely unchanged over time, as it can be challenging to link or connect children's scores across data collection waves for growth purposes when the measures themselves change (see Lloyd, Zumbo, & Siegel, 2009).

In recent years, innovative statistical methods have made it possible to model growth with different scales of measurement (e.g., ordinal scales; see Mehta, Neale, & Flay, 2004) and using ranks outcomes even in the context of changing measures (Bezruczko, Fatani, & Magari, 2016; Lloyd et al., 2009; Moses & Kim, 2012; Stefan & Miclea, 2013), and, importantly, to use these techniques with even small samples. Thus, in the present study, we addressed a major limitation of prior studies seeking to examine the early language skills of poor comprehenders, in that none have examined their language *trajectories* over time, thus limiting our understanding of how children with developmental disorders may grow and change with time (Karmiloff-Smith, 1998). To do so, we obtained and re-analyzed data presented in Justice, Mashburn, and Petscher (2013) in which the language skills of fifth-grade poor comprehenders, poor decoders, and typical readers were compared for mean differences at three points in time (15 months, 36 months, and 54 months). Using these data with a total of three time points with a nearly three and a half year interval between the first and final measurements, we applied multiple group quasi-simplex models as well as multiple group non-parametric linear modeling (NPAR-HLM; Lloyd et al., 2009) in a latent growth framework. The quasi-simplex modeling allowed us to test the longitudinal stability of scores from 15 months to 36 months and 36 months to 54 months and whether the regression coefficients were invariant across the three groups. The use of NPAR-

HLM allowed us to explicitly test the hypothesis that poor comprehenders exhibit a language trajectory over early childhood that not only differentiates them from future typical readers and future poor decoders, but that corresponds to the Matthew Effect.

Methods

Participants

The data in this study were drawn from the database of the NICHD Early Child Care Research Network (NICHD ECCRN, 1993). This national study involved 1,364 children whose mothers gave birth in 1991 in selected U.S. hospitals during 24-hour sampling blocks at 10 separate locations across the country. Efforts were embedded in the sampling procedures to ensure that infants enrolling in the study were relatively healthy: for instance, mothers were required to have no known substance abuse problems and children were required to have no serious medical issues or disabilities. Upon birth, children were followed longitudinally through adolescence to study the effects of various child care experiences on children's development. The resulting sample was relatively diverse (24% of families were ethnic minority status). Additional details on study methodology are readily available (see http://secc.rti.org), and the database is available in the public domain.

For the present study, we investigated the early language skills of 62 children drawn from the larger database w using procedures described by Justice, Mashburn, and Petscher (2013). Selection procedures were as follows. First, we limited the sample to only those children who had completed fifth grade reading Second, of the 991 children in the more limited sample, we identified three groups of readers – poor comprehenders, poor decoders, and typical readers - following procedures discussed in Catts et al. (2006). Note that Catts' prior work on poor comprehenders as well as those of others (e.g., Swanson, Howard, & Saez, 2006) uses the 25th percentile as a cutoff differentiating typical from poor readers. There are certainly arguments that can be made regarding the utility of this normative cutoff, including that it is overly stringent, and we recognize the arbitrary nature of any cut point applied to continuous data; however, for our purposes we adhered our selection to that of the only prior study of poor comprehenders' language skills conducted in the United States (Catts et al., 2008).

For our purposes, poor comprehenders (*n* = 16; 57% boys, 43% girls) were those children who scored at the 25th percentile or lower on PC and at the 40th percentile or higher on LWID at fifth grade. On average, poor comprehenders scored at the 20th percentile rank (*SD* = 5.2) on PC and the 48th percentile (*SD* = 6.9) on LWID. Poor decoders (*n* = 11; 64% boys, 46% girls) were those who scored at the 40th percentile or higher on RC and at the 25th percentile or lower on LWID. The poor decoders scored, on average, at the 48th percentile rank (*SD* = 5.9) on PC and the 21st percentile (*SD* = 3.5) on LWID. Finally, typical readers (*n* = 35; 43% boys, 57% girls) were those who scored between the 40th and 84th percentile on both RC and LWID; of the 349 children who were identified as typical readers, we randomly selected 10% for inclusion in this study. On average, typical readers scored at the 62nd percentile rank (*SD* = 12.5) on PC and the 67th percentile (*SD* = 12) on LWID. Note that at the 5th grade assessment point, the standardized mean difference (Cohen's *d,* bias corrected) in PC scores between poor comprehenders and typical readers was 8.37, and 4.89 between poor comprehenders and poor decoders.

The 62 children were primarily White (74%) and from middle- to upper-SES households (80%) based on household income-to-needs ratios. With the study ascertained at birth in 1991, the fifth-grade assessment time-point used in this study occurred around 2001.

Measures

A comprehensive battery of measures were collected from children from birth into adolescence. Those of relevance to this study are measures of language skill administered prior to school entry, which occurred at 15 months, 36 months, and 54 months, and measures of reading skill collected at fifth grade. Regarding the language measures, 13 different subtests and measures of language skill were administered indirectly or directly to children between birth and school entry. In this study, we selected a language comprehension measure at each assessment wave for consistency in the modality of language being assessed, thus focusing on reception of language versus expression. Each of the three comprehension measures examined children's developing skills in both vocabulary and grammar, including understanding of nouns, verbs, and prepositions (vocabulary) and their construction into phrases, clauses, and sentences (grammar). At 24 months, no measure of comprehension was available; therefore the 24 month time period was not included in our analyses. Information about each of these measures is provided only briefly here, given that codebooks detailing each measure are publicly available.

At 15 months, the *MacArthur Communicative Development Inventory* (CDI; Fenson et al., 1993) was used to assess children's language comprehension skills based on the Phrases Understood subtest. This subtest is completed by children's primary caregivers. There has been extensive study of the utility of the CDI, including a comprehensive evaluation of its measurement properties by Feldman et al (2000) in a study involving 2,156 1- and 2-year-old children. Results of such studies have shown that the CDI subscales – including the Phrases Understood subtest – provide a valid representation of the course of language growth over the toddler years. Feldman and colleagues concluded that parental report is a valid and reliable means for estimating young children's language skills, but argue that it may not be a sensitive means for identifying children who are at-risk for or exhibiting a language disorder. As the CDI is used in this study as a continuous variable characterizing trajectories in language skill over the first two years of life, it is an appropriate means for modeling language growth in these early years. (An overview of psychometric qualities of this tool as aggregated over multiple reports is available at http://www.acf.hhs.gov/programs/opre/ehs/perf_measures/ reports/resources_measuring/res_meas_cdiaa.html.)

At 36 months, the *Reynell Developmental Language Scale* (RDLS; Reynell, 1990) was administered to assess language comprehension using the Verbal Comprehension scale. This standardized assessment was administered during a lab visit by research staff. Independent studies of the psychometric characteristics of the RDLS have supported the validity and reliability claims of the tool's authors (see Udwin & Yule, 1982).

At 54 months, the *Preschool Language Scale-3* (PLS-3; Zimmerman, Steiner, & Pond, 1979) Auditory Comprehension subscale was administered to assess language comprehension. This measure was administered by a researcher during a home visit. Psychometric analyses available from the test developers indicated that it had adequate concurrent validity (e.g., all *r*s > .69 with other standardized measures of language) and reliability estimates.

At fifth grade, children were administered the Woodcock Johnson Test of Achievement (Woodcock & Mather, 1989) Passage Comprehension (PC) and Letter-Word Identification (LWID) subtests, referenced previously. The PC examines children's comprehension by having them read a short passage and identify a key missing word, whereas the LWID examines word recognition by asking children single words of increasing complexity.

Analytical Approach

Specification of the quasi-simplex model was conducted in latent variable framework, whereby the measurement error of the observed variable for each time point was include in the model. A quasi-simplex model is similar to an autoregressive model where the observed measures are regressed onto the immediately previous wave of the measure. For example, with three waves of a language assessment (i.e., *t1, t2, t3*), a quasi-simplex model includes a prediction of the time 3 language assessment from time 2 language, as well as a prediction of time 2 language from time 1 language. The standardized regression coefficients then provide an estimate of stability for the construct assessment between pairs of waves (i.e., time 1 to time 2 and time 2 to time 3). In our specification of the quasi-simplex model, we estimated autoregressive effects using latent variables to account for the measurement error in the observed scores. Because the latent factors were described by single indicators, the factor was identified by fixing the loading to a value of 1.0, and fixing the residual error variance of the observed measure to $(1-\alpha)*\sigma^2$. Once the measurement model was specified, a SEM analysis regressed the latent factor of third measurement occasion (i.e., language comprehension at 54 months; flc54; Figure 1a) on the latent factor of the second measurement occasion (i.e., language comprehension at 36 months; flc36; Figure 1a), which was then regressed on the first measurement occasion (i.e., language comprehension at 15 months; flc15; Figure 1a). In order to test the extent to which the poor comprehenders, poor decoders, and typical readers varied in the autoregressive coefficients, the quasi-simplex model was run in a multiple-group framework. Five sequential models were tested where Model 1 constrained the regression coefficients and residual variances to be equal across all three groups, Model 2 freed the regression constraints for the poor comprehenders, Model 3 added to Model 2 by also freeing the regression constraints for the poor decoders, and Model 4 freed constraints for all subgroups. Because these four models can be viewed as nested, a chi-square difference test was appropriate to evaluate which model provides differential fit from the baseline model.

The NPAR-HLM was subsequently used to examine rank change of language skills among young children who will eventually become poor comprehenders, and to compare their change to those who will become poor decoders as well as those who will be typical readers. NPAR-HLM is a statistical approach that may be used to measure *subgroup change* when measures change over time as children's development is studied, a common occurrence in investigations of such global cognitive domains as language. For instance, measures of very young children's language skills (i.e., those between 12 and 24 months) typically involve parent-report checklists whereas measures of toddlers and preschoolers often involve direct assessment. To address the issues that emerge when using multiple measures over time, even those examining a single construct, NPAR-HLM quantifies changes in relative rank order across time. We should note here that the NPAR-HLM is not specifically a non-parametric analysis, in the way that a Mann-Whitney U test is a non-parametric analog for the parametric *t*-test. Rather, the non-parametric reference in the name pertains to the rescaling of raw data into ranks. As such, it is appropriate to frame this analysis as a hierarchical linear model of rank data; however, in order to provide continuity between the Lloyd et al.'s (2009) explication of the analysis, and our application in this paper, we opt to use the same terminology and refer to this approach to modeling development as a non-parametric approach.

The foundation of this analysis is to provide a mechanism (i.e., a "workable solution") under which one is able to examine change in the presence of changing measures over time. Because different measures of language were not made to be identical, there is difficulty in adjusting the scores to be comparable across measures (i.e., equating) when used over time within a given sample (Kolen & Brennan, 2004); moreover, the sample size in a number of developmental studies may not be large enough to achieve a reliable score transformation among the measures using more advanced measurement techniques, such as item response theory. This is particularly true when studying poor comprehenders, a small subgroup of poor readers. Thus, the advantages of the NPAR-HLM analysis is that it: 1) utilizes the ordinal nature of the data (i.e., children with low scores at a time point have a low rank); 2) does not necessitate common items or tasks over time; 3) helps to bridge the gap between parametric and non-parametric analysis (Conover & Iman, 1981); and 4) provides an initial mechanism for studying change when measures change over time and when one is working with relatively small samples. Mehta, Neale, and Flay (2004) covered a brief history of using growth curves for ordinal data, primarily in the area of adolescent substance abuse, and provided a broad framework for assessing growth in ordinal outcomes when the outcome is the same over time (e.g., rankings of height over time). Work by Lloyd et al. (2009) on the NPAR-HLM extends the logic of Mehta et al. by providing a solution to looking at change when the measures change.

In its simplest form, the NPAR-HLM analysis is conducted by rank transforming the raw data within each time point, and using the ranks as the unit of analysis in the growth model. Inherent to this analysis is that once the raw data are converted, the question of change no longer relates back to the original measures, but is instead framed as a question of change in ranks. Although this may be seen as a limitation of the design, such limitations apply to other non-parametric rank tests such as the Spearman correlation and signed rank tests (Lloyd et al., 2009). Simulation work by Moses and Kim (2012) indicated that the NPAR-HLM approach produces results for measuring change based on ranks comparable to those based on raw, equated scores (i.e., adjusted for differences across groups/forms) relative to parametric methods. Moreover, the authors found that the procedure generates accurate estimates for comparing subgroup change. A noted limitation of using NPAR-HLM when the measures are not identical is that a loss of power could be observed.

Recent evidence also points to a close correspondence between change over time when comparing ordinal scores with residualized change and logit measures. Bezruczko, Fatani, and Magari (2016) found that pre-post change from fall to spring using ordinal measures correlated with logit-based change at $r = .94$ as well as a perfect correlation between ordinal change and the residualized change score $r = 1.00$ using the Woodcock-Johnson III Letter-Word Identification subtask. Such findings do not suggest that the NPAR-HLM approach is a panacea for assessing change over time, especially when interval data that are vertically scaled are available; yet in the absence of common measurement and assessment, the NPAR-HLM may be a viable approach to understanding change over time.

Three assumptions are noted by Lloyd et al. (2009) as inherent to the NPAR-HLM analysis; these must be explored in our dataset before conducting further analysis. First, the scale of the scores from the measures cannot be categorical in nature, as there is not a way to meaningfully convert the original score into a rank. Second, across the waves of data there must be individual differences in change; that is, individuals cannot stay at the same rank across all waves on the measures. If all individuals maintained the same level of rank across the waves, their change score would effectively be zero. As noted by Petscher and Schatschneider (2011), when the data modeled are continuous in nature, this type of trend indicates that individuals are growing, but the reliability of the change score is low because individual differences in growth are not observed. In the context of the NPAR-HLM analysis, heterogeneous growth is of greater importance, as the relative change in scores (i.e., ranks) cannot be related back to an interval metric for comparison of change. By checking the correlation between the initial status rank and the change score, it is possible to describe two types of heterogeneous growth: (a) mastery learning, evidenced by decreasing variances from initial to final testing or (b) fan-spread growth, evidenced by increasing variances from initial to final testing (characterizing the potential presence of a Matthew Effect). A correlation of zero between initial and final testing would provide evidence of individuals maintaining the same rank across time, a violation of the second assumption. Third, a comparable construct must be measured across all waves. Although this assumption may not be specifically tested in many small sample datasets (Lloyd et al., 2009), researchers using the NPAR-HLM approach must make the case that the dimension assessed is commensurate across the different measures. In the present work, although the measures change over time, all are viewed as representing the construct of language, which is generally conceived as a unitary construct (Tomblin & Zhang, 2006). A final component, not discussed directly as an assumption by Lloyd and her colleagues (2009), is related to the completeness of data in one's dataset. Because the original data for this analysis are converted to ranks, the NPAR-HLM may only be conducted with datasets having complete data. When applying a typical growth model, it is plausible to estimate slopes for individuals with missing data points (Singer & Willett, 2003); however, in the presence of the converted data, it is imperative that range of ranks is consistent across all measures. For example, if a sample of 100 participants were administered a measure at time 1, and all participants completed the measure, individuals would be ranked from 1 to 100. At the second time point, if only 90 participants of the original 100 filled out the measure, the rank order would range from 1 to 90. Subsequently, if the lowest scoring individual at both time points was the same individual, their rank would increase in the growth analysis by 10 (i.e., 100 to 90) solely due to missing data. One methodology for controlling the effects of missing data, assuming the data are missing at random or completely at random, is to conduct a multiple imputation of the data prior to the transformation of the raw data into ranks.

An important limitation, acknowledged by Lloyd and colleagues when using NPAR-HLM, is that the method is limited in interpreting model interactions (Sawilowsky, 1990). When one wants to examine differences between groups in growth over time, dummy-code covariates are created, which facilitate a test of mean differences. The structural form of a growth model, whether by traditional multilevel notation or structural equation model notation, defines the test of differences in slopes as an interaction between the dummy-code covariate and the variable representing time. In order to correct for such limitations, we opted to not use a traditional mixed model approach to the analysis, but, rather, we used a multiple-group SEM analysis to test the invariance of growth trajectories across groups. The advantage of conceptualizing the approach as a multiple group problem is that it overcomes interactions as the test of differences between groups in slopes. Rather, it uses parameter constraints to test whether a specific effect should be held constant or freely estimated across groups. By simultaneously fitting individual covariance matrices for the growth model in each group, and forcing a parameter (e.g., the slope) to be the same across the poor decoders, poor comprehenders, and typical readers, we can evaluate how well the model fits compared to one which allows for slopes to be freely estimated for all, or only some, of the groups.

In applying the multiple group latent growth curves, a set of five models was estimated for the language outcomes. Model 1 tested a baseline, fully constrained model which held the initial status, slopes, residual variances, and covariance between intercept and slope equal across all groups; Model 2 freed the initial status and slope means and variances only for poor comprehenders; Model 3 additionally freed initial status and slope means and variances for poor decoders; and Model 4 additionally freed the initial status and slope means and variances for typical readers. Across all model permutations for both the quasi-simplex and NPAR-HLM analyses, a bootstrapping procedure was used for the model coefficients, which has been shown to be advantageous in estimating the standard errors for the effects when sample sizes are small (Efron, 1981). For analyses, due to the small sample size, a *p* –value of .10 was viewed as significant for the purposes to explicating the results.

Results

Prior to conducting the main analyses, we assessed the relative completeness of the data, The range of missing data at various time points was 3.2% to 27.4%, with a mean of 14.7% *(SD* = 10.67%; see Table 1) across all data points. Little's test for data missing completely at random (Little, 1988) was conducted to estimate the extent to which a multiple imputation would produce unbiased estimates of correlations and covariances. The resulting chi-square test [$\chi^2(35) = 37.57$, $p = .352$] indicated that no identifiable pattern existed for the missing data. An examination of the data indicated that scores were not missing due to the observed measures themselves, which is minimal criteria for establishing data are missing at random (Enders, 2010). Subsequently, a multiple imputation with 25 imputations using the expectation maximization technique in SPSS Missing Value Analysis (SPSS, 2012) was conducted, with resulting descriptive statistics reported in Table 1 alongside estimates computed prior to the imputation. All standardized differences between pre- and post-imputed scores were well below the conventional threshold of .20 for a practically important difference (Cohen, 1988).

A secondary descriptive analysis is reported in Table 2, whereby the raw scores for the imputed data were converted to standard scores with a mean of 100 and a standard deviation of 15. The means are reported for the poor comprehenders, poor decoders, and typical reader groups to allow for a comparison of scores across each measure and time point. The results highlighted that poor comprehenders were generally below the mean in language comprehension at each time point relative to their peers. Correlations among the raw data highlighted that a weak association was observed between language comprehension at 15 months and 36 months ($r = .18$), as well as between 15 months and 54 months ($r = .15$); however, a strong correlation was found between 36 and 54 month language comprehension ($r = .67$).

Quasi-Simplex Model

Results from the structural analysis of the quasi-simplex model are reported in Table 3 and Figure 1a. Model 1 for the multiple group analysis, where a completely invariant model was specified across all groups (i.e., residual variances and regression coefficients) as well as the two autoregressive paths were constrained to be equal (i.e., effect of language comprehension at 54 on language comprehension at 36 months and the effect of language comprehension at 36 on language comprehension at 15 months). This model resulted in poor fit to the data [$\chi^2(8) = 29.89$, CFI = .25, TLI = .15, RMSEA = .364, 95% Confidence Interval (.230, .507), p-close < .001]. Model 1 results suggested fit could be improved by freeing the regression of language comprehension at 54 months on language comprehension at 36 months for the poor comprehenders subgroup. By freeing this constraint, model fit significantly improved ($\Delta\chi^2 = 12.39$, $\Delta df = 1$, $p < .001$). Model fit was further improved by relaxing this same regression constraint for the poor decoders in Model 3($\Delta\chi^2 = 8.53$, $\Delta df = 1$, $p = .003$), as well as for the typical readers in Model 4 ($\Delta\chi^2 = 3.95$, $\Delta df = 1$, $p = .025$). As such, we retained Model 4 as providing adequate fit to the data. A depiction of this model [$\chi^2(5) = 3.95$, CFI = 1.00, TLI = 1.00, RMSEA = .000, 95% Confidence Interval (.000, .171), p-close = .586] is shown in Figure 1a, whereby the standardized coefficients from the path diagram for the model shows that the effect of language comprehension at 36 months on language comprehension at 54 months varied across the three subgroups and was strongest for the poor comprehenders (.95), followed by the poor decoders (.91), and the typical readers (.70). Note that the effect of language comprehension at 15 months on 36 months (.29) was found to not vary across subgroup and was weaker than the later autoregressive relation. This finding suggested that greater stability in performance was observed during the latter two assessment periods compared to the first two. That is, children with high language comprehension at 36 months tended to maintain their relative position in the rank ordering of children in this sample at 54 months. Conversely, the relatively weaker coefficient of .29 suggested that greater variability in the rank ordering of children occurred between 15 and the 36 months.

Multiple Group Latent Growth Model (NPAR-HLM Results)

Though the results from the quasi-simplex model demonstrated high individual differences in performance in early language comprehension, along with high stability in later language comprehension, the model did not allow the examination of whether subgroups of children differed in average language comprehension change across the three time points. This shortcoming can be resolved by using the NPAR-HLM procedure. The first assumption of the NPAR-HLM analysis (i.e., that the scores are not categorical) was met. The second assumption for the procedure was examined by specifying a change model which was only conditional on time. There are a number of ways to code time in a traditional growth mode (see Singer & Willett, 2003). Though some may opt to record time as a function of equal intervals over multiple time points (e.g., 0, 1, 2 for three measurement occasions), it is generally more appropriate to code time with greater flexibility. Data collected in this study occurred at unequal intervals, thus, coding time according to the month of assessment will result in a more unbiased estimate of change over time. In our design, we opted to center time on the first wave (i.e., 15 months); therefore, the coding of time in our model was 0 (for 15 months), 21 (elapsed time between 15 and 36 months), and 39 (elapsed time between 15 and 54 months). With this specification of time, the interpretation of the rank change coefficient becomes the average amount of rank change per month.

The unconditional latent growth model estimated the fixed effects (i.e., grand means) of initial status (centered at 15 months) and rank change, as well as the variances for each parameter and the covariance between the two parameters. Upon completion of the imputation, the data were converted from their original metric to rank scores using the RANK command in SPSS (2012), whereby the highest score for each variable was set to a rank of "1", and tied values were assigned a lower rank, rather than the average or high. With the conversion of the imputed raw scores to ranks, it is important to contextualize the ordering and meaning of scores. A rank of 1 represents the top performing individuals, whereas a rank of 62 is associated with the lowest performing individual. Subsequently, when rank change coefficients are estimated, a negative value represents performance improvement (i.e., the more a rank *number* decreases, the higher their performance). For example, a rank of 32 with a mean rank change of -.20 would indicate that, on average, children's rank was decreasing by .20 units. Over the course of five months, the average rank would change from 32 to 31 [i.e., 32 + (-.20*5)], thus showing that the average child's performance *improves* with time. With both empirically-based assumptions tested, the results suggested that the NPAR-HLM was appropriate to apply to the imputed, rank-transformed scores. An examination of the correlation among student ranks scores yielded nearly identical estimates as those reported on the raw data [i.e., 15 months and 36 months (r = .18), 15 months and 54 months (r = .21), and 36 and 54 month language comprehension (r = .67)].

The unconditional model (Figure 1b) shows that the mean rank across all children was approximately 32. Because the data were centered at 15 months, this value is reflective of the mean rank at the first assessment point. Although this was statistically significant ($p < .001$), it simply means that the mean rank was significantly different from zero. The associated mean rank change language comprehension was 0.02. Both of these preliminary findings are in keeping with expected results from an unconditional latent growth model of rank data, in that the mean initial status is expected to be the median ranks, and the mean change is expected to be zero. Although the mean rank change was not significant, the variance of rank change *was* for language comprehension (i.e., 0.15, Figure 2) signifying that children varied in the amount of change that occurred over the waves of assessment. By modeling children's language differences as a multiple indicator in the model (i.e., typical reader, poor comprehender, poor decoder), it was possible to test for differences in performance across the groups in both initial status and rank change. Of particular importance in Figure 1b is that the covariance between intercept and slope was -2.06; as a standardized correlation, the estimate is -.42 meaning that there is a moderate, negative correlation between initial language comprehension and change over time. In conventional growth models, the correlation between initial status (when centered at time 1) and change is a reflection of the functional form of growth (i.e., mastery learning, parallel growth, or fan-spread growth/Matthew effects). The conceptual interpretation of the negative covariance/correlation here is that individuals with a higher *score* change less over and individuals with a lower *score* change more. Because these data reflect ranks, there is an inverse property at play; that is, lower scores (e.g., 1) indicate higher rank and higher scores (e.g., 50) indicate lower rank. These data then convey that even though individuals with lower *scores* change more *those with lower scores have higher ranks and vice versa*. The implication of this interpretation is that the negative covariance actually reflects that individuals who are doing better (i.e., higher rank) change more (i.e., continue to lower score/increase their rank) compared to individuals with a higher score (i.e., lower rank) who change less. In this way, the data from the unconditional model indicate that students with better ranks improve their ranks more than poor performing students, an indication of the potential presence of a Matthew Effect. As a second test of the possible presence of Matthew Effects, the growth model of Figure 1 was refit so that the intercept was centered at 54 months, rather than 15 months. By comparing the variances from the model where intercepts were centered at 15 months versus 54 months, it was possible to evaluate whether variances were increasing, decreasing, or staying approximately the same. These separate variance estimates are reflected in Figure 1 where the variance in scores at 54 months (i.e., 228.72, $p < .05$) was greater than that observed at 15 months (160.40, $p < .01$). The confluence of the increasing variances and nature of the correlation between intercept and slope both meet the criteria described by Zumbo (1999) of the possibility of a Matthew Effect.

When testing the multiple group models with constraints for language comprehension, Table 3 summarizes the χ^2 values for the models where initial status and rank change means and variances were: 1) constrained to be equal across all subgroups ($\chi^2 = 44.97$, Model 1), 2) freed for the poor comprehenders while the poor decoders and typical readers were constrained to be equal ($\chi^2 = 30.15$, Model 2), 3) additionally freed for the poor decoders ($\chi^2 = 18.79$; Model 3), and 4) additionally freed for the typical readers ($\chi^2 = 15.27$; Model 4). A comparison of each nested model showed that freeing the means and variances of the initial status and rank change generally led to significantly improved model fit. As a further improvement to Model 4, a final model (i.e., Model 5) freely estimated a residual variance for the typical readers. Fit for this model was acceptable [$\chi^2(11) = 11.07$, CFI = .99, TLI = .99, RMSEA = .017, 95% Confidence Interval (.000, .232), *p*-close = .484].

Figure 2 presents the results for the Model 5 language comprehension outcome. For both the means and the variances, the first coefficient reported in Figure 2 represents the value for the poor comprehenders, the second coefficient is for the poor decoders, and the third coefficient is for the typical readers. The intercept in this model represents the mean rank at 15 months for children who would become poor comprehenders, and was estimated at approximately 38, and suggested that when compared to their peers, children who would become poor comprehenders started in the middle part of the distribution of children. Figure 3 plots mean rank at 15 months for the poor comprehenders along with the change in rank over time. Recalling that for the rank scores, lower values reflect better performance, and higher values mean reflect performance, Figure 2 shows that the mean rank change per month for the poor comprehenders was 0.26 for language comprehension. Note again that when interpreting the rank change means, ordinal ranks are being modeled over time; *thus, rank change is inversely related to progress*, such that an average *negative* change indicates that rank/performance is improving, whereas an average *positive* change indicates that rank/performance is decreasing relative to peers. Figure 3 shows that poor comprehenders who started with a mean rank of 38 ended with a mean rank of approximately 50 by age 54 months.

Children who would become poor decoders had mean rank scores of 41 for language comprehension, which approximated the poor comprehenders (see Figures 2 and 3). The poor decoders' growth rate was -0.49 units per month, and was statistically different from typical readers and poor comprehenders. The negative slope mathematically means that the average rank became smaller over time; conceptually, a smaller rank indicates that performance is improving over time (i.e., their rank moved from a larger number/lower rank to a smaller number/higher rank; see Figure 3). Children who would become typical readers had an approximate mean initial rank of 28 in language comprehension, indicating they started off with higher language ability than both future poor comprehenders and poor decoders. Pertaining to their rank change, typical readers' rank decreased over time, changing an average of -.04 ranks per month, indicating that typical readers were: 1) improving their rank over time compared to the poor comprehenders whose rank became worse over time and 2) improved their rank over time but at a rate less than that of the poor decoders (Figure 3).

Because the statistical significance of the poor decoder slope was at the level of $p < .10$, we opted to compute standardized effect sizes for each group to characterize the magnitude of change in rank over time. This process was done by taking the difference between their 15 month and 54 month rank and dividing by the standard deviation of 15 months within group. Results suggested that the magnitude of the poor comprehender change was moderate (i.e., d = -0.64) to characterize their downward performance over time. Conversely, the positive, moderate effect for rank change with the poor decoders (d = 0.58) reflects the improved ranking over time. The lack of change in rank for the typical readers was reflected by the small effect size (d = -0.10). Figure 3 demonstrates the rate of change for the three groups in language comprehension, and highlights the trend of convergence in rank for language comprehension between the future typical readers and future poor decoders. In addition, Figure 3 shows that the poor comprehenders lag increasingly behind their peers from 15 months forward, consistent with the idea of a Matthew Effect.

Discussion

The objective of this study was to examine the nature of early language change, from 15 to 54 months, for children who would become poor comprehenders at fifth grade in comparison to children who would become typical readers or poor decoders. Results of this study, which used two approaches to modeling early language trajectories when measures of early language skill change over time, showed that children who are poor comprehenders at fifth grade fall increasingly behind typical readers and poor readers in their language skills over time. That is, their language skills worsen over time in relation to other types of readers. The results of this study provide convincing evidence that poor comprehenders exhibit an underlying weak language system, a weakness that becomes increasingly apparent as language demands increase as children enter into school and must use their linguistic resources to read for meaning. Further, the results suggest that poor reading comprehension may be conceptualized as a symptom of an underlying developmental language impairment that is present quite early in life.

For some time, experts have surmised that the reading difficulties of poor comprehenders likely reflect a developmental language disorder which was present well before the advent of formal reading instruction (see Hulme & Snowling, 2009). Although studies have shown that the language skills of children who are poor comprehenders differ substantially from those of typical readers and poor decoders at 4 and 5 years of age (Justice, Mashburn, & Petscher, 2013; Nation et al., 2010), the actual trajectory of very early language growth among poor comprehenders has not been empirically assessed. We might anticipate, if children who go on to be poor comprehenders do indeed exhibit a developmental language impairment during early childhood, that their trajectory of early language acquisition would look different than that of other groups of readers. Moreover, research reported by Catts et al (2006) and Justice, Mashburn, and Petscher (2013) suggested that the Matthew Effect might characterize the language growth of future poor comprehenders in relation to their peers: the Matthew Effect is present when initial skill levels predict growth over time, with those with lower initial skills falling increasingly behind their peers consistent with a gap-widening effect (or fan-spread change). Presence of a Matthew Effect might suggest that a defining feature of language acquisition among poor comprehenders is its developmental course over time.

In examining the early language trajectories of poor comprehenders, two contributions of the present study are highlighted. First, an important and unique contribution of the present study is that we used novel statistical methods to model change in language skills over a nearly 40-month period when the measures used to assess language changed themselves. This technique allowed us to extend prior results reported in Justice et al (2014) and Catts et al (2006), which suggested that the language skills of poor comprehenders may decline in relation to other children's language skills over time. A commonality of both such studies is that the measures used to assess children's language skills changed over time; thus, the studies could not definitively measure growth trajectories. To address the issues that emerge when using multiple measures over time, NPAR-HLM involves quantifying change in relative rank order over time by rescaling raw data into ranks. The application of methods that use ordinal units for understanding change are relatively new (Bezruczko, Fatani, & Magari, 2016; Humphry & McGrane, 2015; Lloyd, 2010; Lloyd & Hertzman, 2009; Lloyd, Zumbo, & Siegel, 2009; Mehta, McNeale, & Flay, 2004) yet emerging evidence has demonstrated that ordinal and linear gains demonstrate strong consistency (Bezruczko, Fatani, & Magari, 2016). By applying this approach to the early language development of children who become poor comprehenders at fifth grade relative to two other groups of readers (future poor decoders and future typical readers), the present results showed that poor comprehenders' language comprehension lags increasingly behind that of other children from 15 months to 54 months, becoming increasingly divergent from that of other children.

Second, an additional contribution is demonstration of the obvious divergence in language comprehension that characterized the rank change of poor comprehenders compared to poor decoders from 15 to 54 months. The growth models that poor comprehenders and poor decoders had similar language skills at 15 months but that these skills diverged substantially with time, with the poor comprehenders' skills falling further behind those of the poor decoders. This finding suggests that identification of future poor comprehenders – a significant concern given that many of these children are not identified in the early grades since their decoding skills progress normally (Nation et al., 2004) – optimally involves examination of language trajectories rather than static assessment at a given point in time. The findings presented here converge with those of Catts and colleagues (2006), which examined the language skills of eighth-grade poor comprehenders, poor decoders, and typical readers at kindergarten, second, and fourth grades. Their work also suggested that the language skills of poor comprehenders lagged substantially over time in relation to other children. In fact, on an index of vocabulary skills, the skills of poor comprehenders declined approximately one-third of a standard deviation from kindergarten to fourth grade; by comparison, the vocabulary skills of poor decoders increased modestly, about one-tenth of a standard deviation, whereas the vocabulary skills of typical readers stayed stable over time.

An important conclusion of the present study is finding of a Matthew Effect to characterize the early language trajectories of poor comprehenders compared to other children. This finding was not unexpected, given that we re-analyzed data presented previously from these same children (Justice et al., 2014), in which comparisons using analyses of variance at each time point (15, 36, 54 months) showed that significant and meaningful differences in language skills between poor comprehenders and the other two reader groups were most striking at 54 months compared to the prior three points. This prior study did not determine, however, whether the language trajectory of children who would become poor comprehenders differed in any fundamental way from that of other groups of poor readers. That the early language skills of future poor comprehenders do appear to lag over time in relation to other children may be an important characteristic of this type of reading disability and the language difficulties these children experience. Obviously, we must consider why such deterioration occurs and, in the future, explore how it might be prevented or remediated. Perhaps the most likely explanation is that in the early childhood years, language skills are undergoing rapid advancements.

As noted in the introduction, a notable early accomplishment in language acquisition is combining words to form multi-word combinations, which often happens between 18 and 24 months. From this period into the later preschool years, children use increasingly complex grammatical forms and acquire an impressively large number of words (Turnbull & Justice, 2016). An important hallmark of language acquisition in the preschool years is the ability to integrate component language skills (e.g., vocabulary, grammar, pragmatics) into more complex discourse structures (e.g., conversations, narratives; see Brown, 1973). For children who have underlying processing limitations that negatively affect language ability, these negative impacts may not be apparent until children are required to use language at these more complex discourse levels. For instance, for children who have developmental language disorders, their abilities to integrate component language skills become particularly apparent in narrative tasks. These children produce stories at three and four years of age that contain fewer words, shorter sentences, and less connected story components compared to their typical peers (Kaderavek &

Sulzby, 2000). The use of more complex tasks to assess children's language skills over time, as occurred in this study, and the inability of children with relatively weak language skills to navigate such tasks, manifests itself in a pattern of deterioration. Longitudinal assessment of the reading trajectories of children with developmental language impairment from kindergarten to fifth grade also show evidence of the Matthew Effect, with these children falling increasingly behind non-impaired peers in their reading skills with time. Presumably this effect occurs due to the increasingly demanding contexts in which they must use their language skills to process oral and written discourse, such as comprehending expository texts. Because children with poor language skills – whether they have developmental language impairments or are poor comprehenders (which may be considered a variant of the former; see Hulme & Snowling, 2009) – may have difficulty deploying their language skills within these increasingly demanding contexts, they have become less motivated than other children to do so and in turn receive less practice. This circumstance can result in achievement lags over time in comparison to other children who are motivated to deploy their language skills in such contexts and do so often.

An interesting finding of the present work that also warrants note concerns the early language trajectories of children who would go on to become poor decoders at grade 5. At 15 months, the future poor decoders showed language comprehension skills that were the lowest among the three subgroups in the study (M = 94.6 vs. 100 for future poor comprehenders and 101.7 for future typical readers). Results available within a recent meta-analysis also showed that children with decoding-specific reading problems had poorer language skills at infancy and toddlerhood than children in a control group (Snowling & Melby-Lervåg, 2016). The present results further show, however, that with time the future poor decoders significantly gained in their language comprehension to become similar to typical readers and disambiguated from poor comprehenders, yet their initially low language comprehension skills are an interesting finding. As poor comprehenders have underlying difficulties with the phonological domain of language (Ziegler & Goswami, 2005), their early lags in language comprehension may reflect the important role of early phonological skills to lexical and syntactic development (Christophe, Millotte, Bernal, & Lidz, 2008). That is, some theories of early language acquisition propose that phonological skills serve to bootstrap syntactic and lexical development, which would contribute to a general lag in early language development for children with core phonological weaknesses. With time, however, future poor decoders appear to "catch up" with future typical readers in their language comprehension skills, as phonological skill plays a less prominent role to ongoing linguistic development. Nonetheless, the underlying deficits in phonological skill are not resolved, and subsequently serve to undermine the poor decoders' ability to develop strong word-recognition skills. The accelerated trajectory of language growth seen for future poor decoders across the toddler and preschool years corresponds to what Scarborough and Dobrich (1990) called an "illusory recovery" so as to explain how children can appear to overcome early language difficulties and then at a later time exhibit reading problems. The present study suggests the need for future research to carefully explore the early language skills of poor decoders, to determine whether early linguistic trajectories may have a prognostic role in early identification of these children as well as future poor comprehenders.

Findings such as those presented here, coupled with studies of the positive predictive relations between language skills and reading achievement among typical readers (e.g., NICHD ECCRN, 2005), have been interpreted by some as indicating that deficits in language abilities are a causal determinant of the reading difficulties experienced by poor comprehenders (see Hulme & Snowling, 2009). Although there has been a large volume of research on the causal determinants of reading difficulties experienced by poor decoders, permitting a strong claim to be made that poor decoding skills arise from core phonological processing deficits (see Vellutino, Fletcher, Snowling, & Scanlon, 2004, for review), research on the causal determinants of poor comprehension has been much more limited.

We acknowledge several limitations pertaining to this work. First, although the analysis of ordinal data, especially those pertaining to Likert scales and ordered categories, are found in the literature, there are relatively few instances of the NPAR-HLM for studies where the measures change and/or using rank data as the outcome. Because the NPAR-HLM and analysis of rank data is relatively new, with the first published paper appearing in 2009, it is imperative that this technique is studied further through new and extended simulations such as those done by Moses and Kim (2012). As other data become available, it is important to replicate and extend the work here to evaluate the inferences put forth. Connected to this limitation, rank data exist as a special case of ordinal data such that while the latter affords for changes in the distributions over time (i.e., the unconditional mean and variance change over time), the former are fixed (i.e., the unconditional mean and variance are unchanged for ranks over time). An important implication is that estimation or ordinal with appropriate link functions should be considered. In the present study, maximum likelihood was used for the data. Depending on the maximum number of ranks, one should be cognizant of the extent to which the rank data are treated as categorical or continuous. In some cases, the number of categories may exceed that which is allowable for treating a categorical outcome as categorical (e.g., Mplus software only allows for 10 ordered categories). Rhemtulla, Brosseau-Liard, and Savalei (2012) provide useful guidance on when categorical variables may be treated as continuous.

Second, related to sample size, the sample of sixty-two may be viewed as small, and there are two implications for a small sample worth unpacking. As it pertains to hypothesis testing, the lack of statistical significance between subgroups may be less due to the lack of a true score relation and rather due to a lack of power for detecting small effects. Further, it is possible that standard errors for a small sample are biased, and thus that any effects observed may not generalize with a larger sample. To guard against small sample bias, the bootstrap was used so that tests of significance were evaluated based on the multiple samples that occur with the bootstrap. This procedure makes the observed effects for differences in rank slope less biased due to small samples.

Third, in the absence of common measurement over time, it is important to contextualize findings such that changes in rank order may not provide the same changes as when common measurement is used to examine longitudinal trajectories in develpment. Though preliminary evidence by Bezruczko, Fatani, and Magari (2016) found consistency between rank and interval-based measurement, replication is needed in this area of research.

An important future goal of research on poor comprehension is to use complementary research methods to further assess the causal relations between oral language skills and comprehension failure, such as treatment studies (Clarke et al., 2010). Treatment studies have provided a causally interpretable means for identifying the causal relations between phonological processing skills and decoding problems: by randomly assigning poor decoders to receive training in phonological processing skills and showing that word recognition improved as a result (as compared to poor decoders who did not receive such training), the causal relations between phonological processing and word recognition could be firmly established (see Vellutino et al., 2004). Such work has been highly influential to the development of practices that can prevent and/or remediate the reading difficulties of poor decoders, and will be important for future research further examining the causal relations between language skills and reading-comprehension difficulties.

References

Aaron, P. G., Joshi, M., & Williams, K. A. (1999). Not all reading disabilities are alike. Journal of Learning Disabilities, 32(2), 120-137.

Adams, M. J. (2010, winter). Advancing our students' language and literacy: The challenge of complex texts. American Educator, 3-11.

Bast, J., & Reitsma, P. (1997). Matthew effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research, 32,* 135-167.

Bezruczko, N., Fatani, S. S., & Magari, N. (2016). Three Tales of Change. SAGE Open, 6(3), 2158244016659905.

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. International Journal of Language & Communication Disorders, 49(4), 381-415.

Bloom, R. (1973). A first language: The early stages. Cambridge, MA: Harvard University Press.

Catts, H. W., Adlof, S. M. & Ellis Weismer, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. Journal of Speech, Language, and Hearing Research, 49, 278-293.

Catts, H. W., Adlof, S., Hogan, T. P., & Ellis Weismer, S. (2005). Dyslexia and specific language impairment: Same or different developmental disorder? Journal of Speech, Language, and Hearing Research, 48, 1378-1396.

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25-40). Mahwah, NJ: Lawrence Erlbaum.

Catts, H. W., Herrera, S., Nielsen, D. C., & Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. Reading and Writing, 28(9), 1407-1425.Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. Language and speech, 51(1-2), 61-75.

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating chidren's reading-comprehension difficulties: A randomized controlled trial. Psychological Science *21,* 1106-1116.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

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.Duff, D., Tomblin, J. B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a Matthew Effect. Journal of Speech, Language, and Hearing Research, 58(3), 853-864.

Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap, and other methods. *Biometrika, 68,* 589-599.

Ellis, E. M., Borovsky, A., Elman, J. L., & Evans, J. L. (2015). Novel word learning: An eye-tracking study. Are 18-month-old late talkers really different from their typical peers?. Journal of communication disorders, 58, 143-157.

Enders, C.K. (2010). *Applied missing data analysis.* New York, NY: Guilford Press.

Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, J., et al. (1993). *The MacArthur communicative development inventories: User's guide and technical manual*. San Diego, CA: Singular.

Givon, T. (2013). Function, structure and language acquisition. The crosslinguistic study of language acquisition, 2, 1005-1028.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. Remedial and special education, 7(1), 6-10.

Hoover, W., & Gough, P. (1990). The simple view of reading. Reading and Writing, 2, 127-160.

Humphry, S. M., & McGrane, J. A. (2015). Equating a large-scale writing assessment using pairwise comparisons of performances. The Australian Educational Researcher, 42(4), 443-460.

Justice, L., Mashburn, A., & Petscher, Y. (2013). Very early language skills of fifth-grade poor comprehenders. Journal of research in reading, 36(2), 172-185.

Kaderavek, J.N. & Sulzby, E. (2000). Narrative production by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 34-49.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*(10), 389-398. doi:10.1016/S1364-6613(98)01230-3.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices.* New York: Springer.

Language and Reading Research Consortium. (2015). Learning to read: should we keep things simple?. Reading research quarterly, 50(2), 151-169.

Language and Reading Research Consortium (In Press). Oral language and listening comprehension: Same or different constructs? Journal of Speech, Language, and Hearing Research.

Law, J., Tomblin, J. B., & Zhang, X. (2008). Characterizing the growth trajectories of language-impaired children between 7-11 years of age. *Journal of Speech, Language and Hearing Research, 51,* 739-749.

Leonard, L. B. (2000). *Children with specific language impairment.* Cambridge, MA: MIT Press.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical Association, 83(404), 1198-1202.

Lloyd, J. E., & Hertzman, C. (2009). From kindergarten readiness to fourth-grade assessment: Longitudinal analysis with linked population data. Social Science & Medicine, 68(1), 111-123.

Lloyd, J. E. V., Zumbo, B. D., & Siegel, L. S. (2009). When measures change over time: A workable solution for analysing change and growth across multiple waves. *Journal of Educational Research and Policy Studies, 9*, 81-100.

McKean, C., Mensah, F. K., Eadie, P., Bavin, E. L., Bretherton, L., Cini, E., & Reilly, S. (2015). Levers for language growth: characteristics and predictors of language trajectories between 4 and 7 years. PloS one, 10(8), e0134251.

Mehta, P. D., Neale, M. C., & Flay, B. R. (2004). Squeezing interval change from ordinal panel data: latent growth curves with ordinal outcomes. Psychological Methods, 9(3), 301.

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. (2014). Declarative and procedural memory as individual differences in second language acquisition. Bilingualism: Language and Cognition, 17(01), 56-72.

Moses, T., & Kim, S. (2012). Evaluating ranking strategies in assessing change when the measures differ across time. Educational and Psychological Measurement. Doi 10.1177/0013164411405651.

Murphy, K.A., Language and Reading Research Consortium, & Farquharson, K. (2016). Investigating profiles of lexical quality in preschool and their contribution to first grade reading. Reading and Writing, 29, 1745-1770.

Muthen, L.K., & Muthen, B.O. (1998-2012). *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthen & Muthen.

Nation, K. (2005). Connections between reading and language in children with poor reading comprehension. In H. Catts & A. Kamhi (Eds.), Connections between reading and language (pp. 41-54). Erlbaum.

Nation, K., Clarke, P., & Snowling, M. J. (2002). General cognitive ability in children with reading comprehension difficulties. *British Journal of Educational Psychology, 72,* 549-560.

Nation, K., Clarke, P., & Marshall, C.M., & Durand, M. (2004). Hidden language impairments in children : Parrellels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research, 47*, 199-211.

Nation, K., Cocksey, J., Taylor, J. S. H., & Bishop, D. V. M. (2010). A longitudinal investigation of the early reading and language skills in children with poor reading comprehension. Journal of Child Psychology and Psychiatry.

National Early Literacy Panel. (2008). Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Institute for Literacy. Available at http://www.nifl.gov/earlychildhood/NELP/NELPreport.html. NICHD Early Child Care

Research Network. (1993). *The NICHD Study of Early Child Care: A comprehensive longitudinal study of young children's lives.* Rockville, MD:Author. (ERIC Document Reproduction Service No. ED3530870)

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.

Newcomer, P., & Hammill, D. (1988). Test of Language Development-Primary 2nd Edition. Austin, TX: PRO-ED.

Petscher, Y., & Schatschneider, C. (2011). A simulation study on the performance of the simple difference and covariance-adjusted scores in randomized experimental designs. Journal of Educational Measurement, 48, 31-43.

Reynell, J. (1990). *Reynell Developmental Language Scale*. Los Angeles: Western Psychological Services.

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. Psychological methods, 17(3), 354.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1412-1431.

SAS Institute (2009). *Base SAS 9.2 Procedures Guide.* Cary, NC: SAS Institute.

Scarborough, H. S., & Dobrich, W. (1990). Development of children with early language delay. Journal of Speech, Language, and Hearing Research, 33(1), 70-83.

Schuele, C. M., & Dykes, J. C. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment. Clinical linguistics & phonetics, 19(4), 295-318.

Semel, E., Wiig, E. H., & Secord, W. A. (1995). *Clinical evaluation of language fundamentals third edition.* San Antonio, TX: The Psychological Corporation.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and even occurrence.* New York: Oxford University Press.

Snowling, N., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review *Psychological Bulletin, 142*, 498-545.

SPSS (2012). *SPSS Base 18.0 for Window's User's Guide*. SPSS Inc., Chicago IL.

Ştefan, C. A., & Miclea, M. (2013). Effects of a multifocused prevention program on preschool children's competencies and behavior problems. Psychology in the Schools, 50(4), 382-402.

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.

Stothard, S. E., & Hulme, C. (1995). A comparison of phonological skills in children with reading comprehension difficulties and children with decoding difficulties. *Journal of Child Psychology and Psychiatry, 36,* 399-408. doi: 10.1111/j.1469-7610.1995.tb01298.x

Tambyraja, S. R., Schmitt, M. B., Farquharson, K., & Justice, L. M. (2015). Stability of language and literacy profiles of children with language impairment in the public schools. Journal of Speech, Language, and Hearing Research, 58(4), 1167-1181.

Turnbull, K. L. P., & Justice, L. M. (2016). Language development from theory to practice. Pearson.

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. Developmental Science, 11, 84-97.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*, 2-40.

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

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. Psychological bulletin, 131(1), 3.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1979). *Preschool Language Scale*. Columbus, OH: Charles Merrill.

Zumbo, B.D. (1999). The simple difference score as an inherently poor measure of change: Some reality, much mythology. In Bruce Thompson (ed.). *Advances in Social Science Methodology,* Volume 5 (pp. 269-304). Greenwich, CT: JAI Press.

Table 1

Descriptive Statistics Pre- and Post-Imputation

	Pre-Imputation				Post-Imputation (N = 62)				
Measure	N	Mean	SD	% Missing	Mean	SD	25th Quartile	50th Quartile	75th Quartile
15 month LC	45	38.67	29.02	27.4	38.22	25.3	15.00	35.58	58.20
36 month LC	60	97.28	15.62	3.2	96.93	15.5	84.00	96.00	106.00
54 month LC	57	44.25	35.25	8.1	43.62	34.07	8.00	44.89	73.00

Note. LC = Language Comprehension.

Table 2

Standard Scores for Each Group of Readers

				Quartile Values		
Group	Measure	Mean	Std. Dev.	Q25	Q50	Q75
Poor Comprehenders	15 month LC	100.00	15.81	85.21	99.27	110.57
	36 month LC	89.09	12.43	80.80	86.60	92.39
	54 month LC	89.58	9.39	81.98	84.13	95.24
Poor Decoders	15 month LC	94.60	16.09	80.74	92.42	96.31
	36 month LC	101.00	13.93	88.53	96.26	109.19
	54 month LC	98.21	14.96	83.25	100.48	107.55
Typical Readers	15 month LC	101.69	14.31	91.08	99.89	108.33
	36 month LC	104.67	14.13	92.39	104.95	112.68
	54 month LC	105.33	14.75	89.88	100.48	117.72

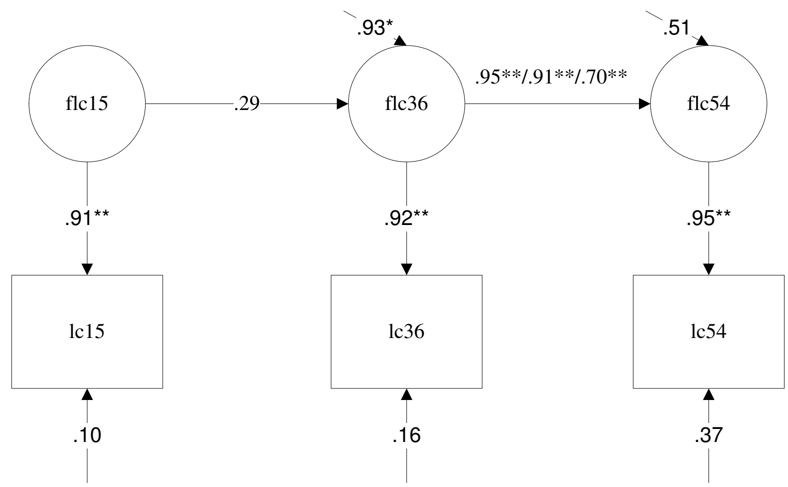
Note. LC = Language Comprehension.

Table 3

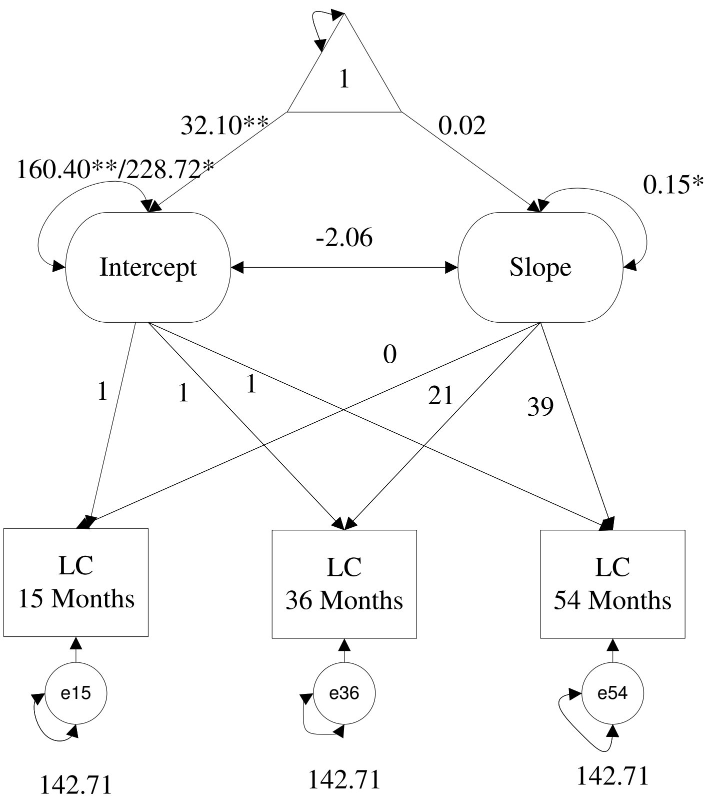
Chi-Square Test Results from Multiple Group Latent Growth Model Comparison

Analysis	Outcome and Model	χ^2	df	$\Delta\chi^2$	Δdf	p-value
Quasi-simplex	Model 1	29.89	8			
	Model 2	17.50	7	12.39	1	<.001
	Model 3	8.97	6	8.53	1	.003
	Model 4	3.95	5	3.95	1	.025
NPAR-HLM	Model 1	44.97	21			
	Model 2	27.05	17	14.82	1	.005
	Model 3	18.79	13	10.36	1	.034
	Model 4	15.27	12	3.52	1	.060
	Model 5	11.07	11	4.20	1	.040

Note. For the Quasi-simplex analysis: Model 1 = fully constrained regression; Model 2 = freed regression for Poor Comprehenders; Model 3 = freed regression for Poor Comprehenders and Poor Decoders; Model 4 = freed regression for all groups. For the NPAR-HLM: Model 1 = fully constrained means for intercepts and slopes, residual variances, and covariance; Model 2 = freed intercept and slope means for Poor Comprehenders; Model 3 = freed intercept and slope means for Poor Comprehenders and Poor Decoders; Model 4 = freed intercept and slope means for all groups; Model 5 = Model 4 + freely estimated residual variance for Typical Readers.



(a)



(b)

Figure 1. Results from a) the standardized quasi-simplex model for the full sample (Model 1) and b) Unconditional latent growth model for language comprehension. For Figure 1a, the three standardized values for the regression of flc54 on flc36 reflects the coefficients for the Poor Comprehders, Poor Decoders, and Typical Readers, respectively. For Figure 1b, the first value for the variance of the intercept reflects centering at 15 months. The second value for the variance of the intercept reflects centering at 54 months. $*p < .05$, $**p < .01$. The direct lines from the triangle are indicative of the latent intercept and slope means.

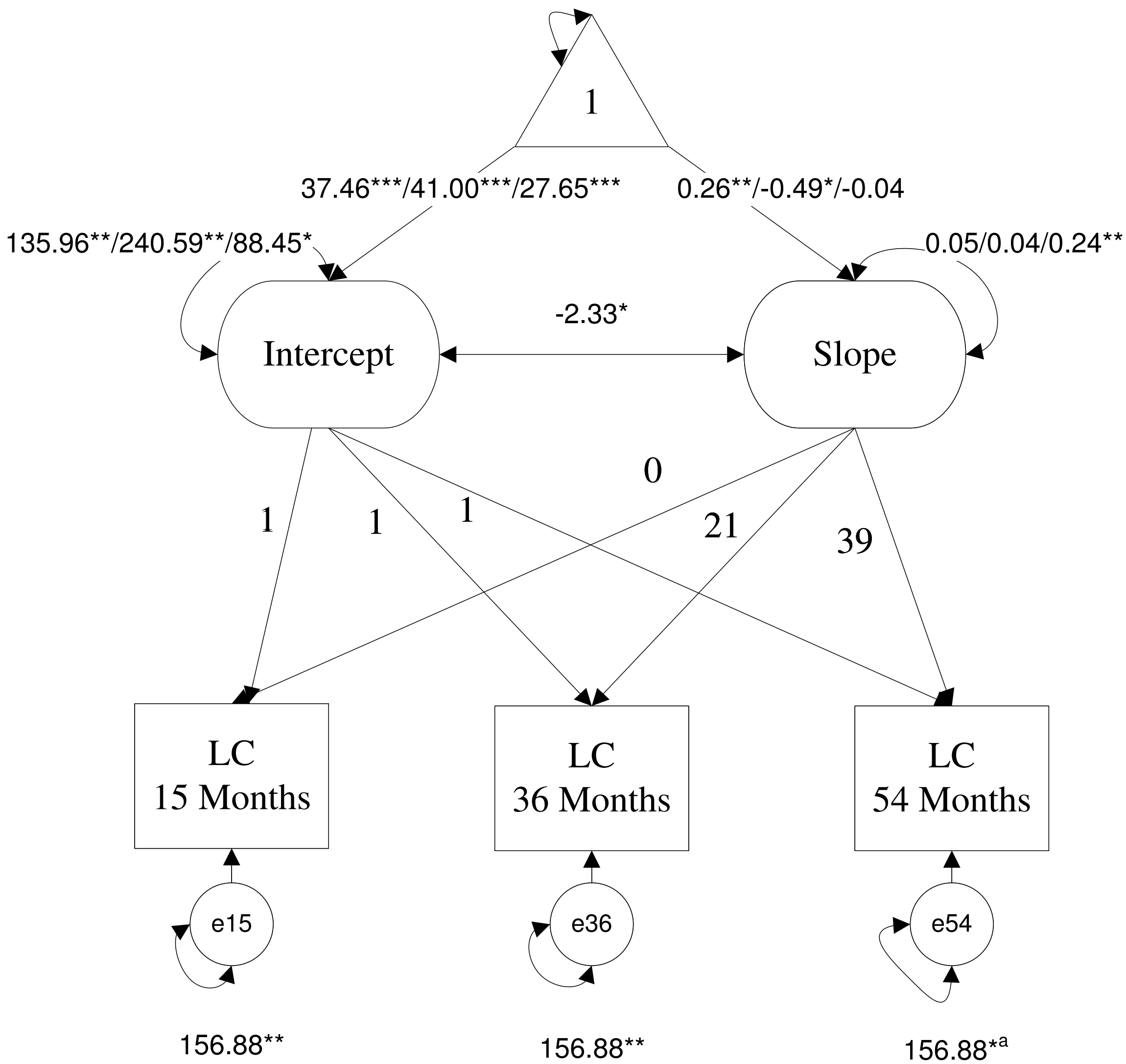


Figure 2. Multiple group latent growth model for language comprehension (Model 5). The first coefficient for the means and variances represents the Poor Comprehenders, the second coefficient represents the Poor Decoders, the third coefficient represents the Typical Readers. *$p < .10$, **$p < .05$, ***$p < .01$. [a]The residual variance for Typical Readers at 54 months was fixed at 0 due to a negative, but non-signifcant value. The direct lines from the triangle are indicative of the latent intercept and slope means by subgroup.

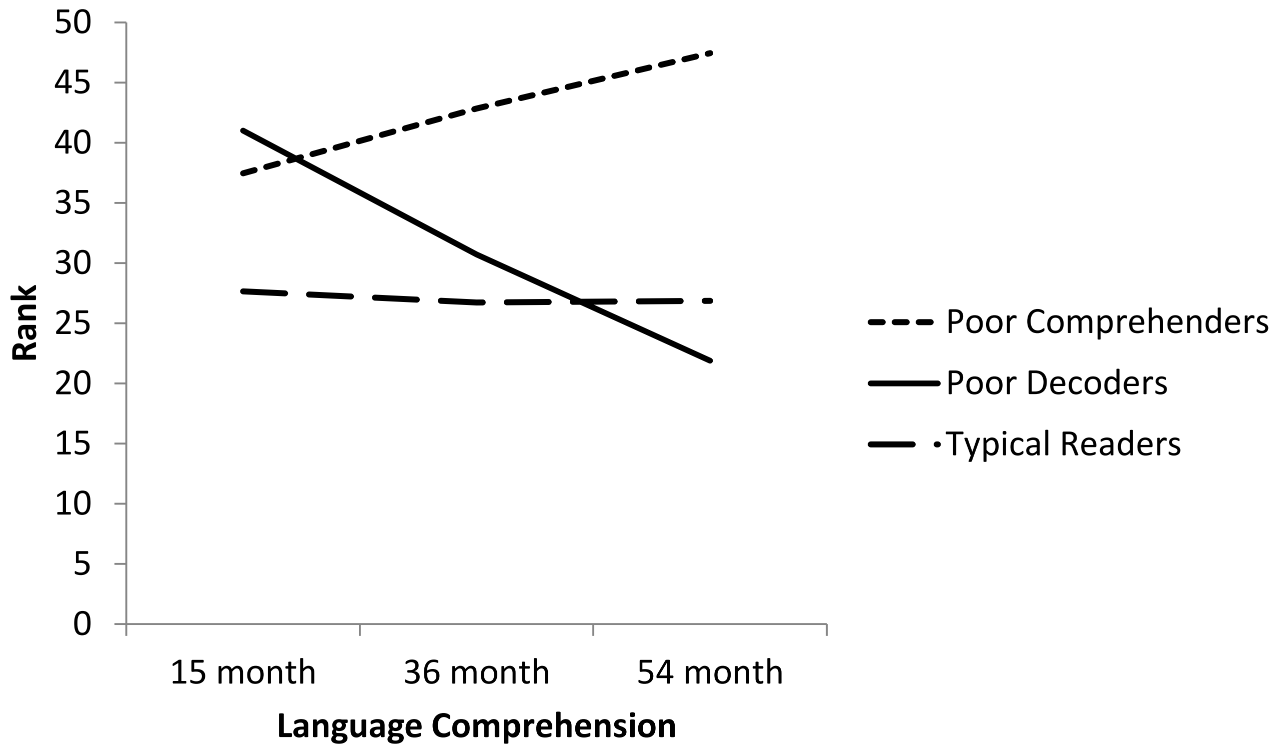


Figure 3. Predicted language comprehension rank trajectories for the Poor Comprehenders (dotted line), Poor Decoders (solid line), and Typical Readers (dashed line).