Published in final edited form as:

J Learn Disabil. 2018 ; 51(6): 565–577. doi:10.1177/0022219417719887.

Examining Predictive Validity of Oral Reading Fluency Slope in Upper Elementary Grades using Quantile Regression

Eunsoo Cho, Ph.D.,

The Meadows Center for Preventing Educational Risk, University of Texas at Austin, 1912 Speedway D4900, Austin, TX 78712, eunsoocho@austin.utexas.edu

Philip Capin,

The Meadows Center for Preventing Educational Risk, University of Texas at Austin, 1912 Speedway D4900, Austin, TX 78712, pcapin@utexas.edu

Greg Roberts, Ph.D., and

The Meadows Center for Preventing Educational Risk, University of Texas at Austin, 1912 Speedway D4900, Austin, TX 78712, gregroberts@austin.utexas.edu

Sharon Vaughn, Ph.D.

The Meadows Center for Preventing Educational Risk, University of Texas at Austin, 1912 Speedway D4900, Austin, TX 78712, srvaughn@austin.utexas.edu

Abstract

Within multitiered instructional delivery models, progress monitoring is a key mechanism for determining whether a child demonstrates an adequate response to instruction. One measure commonly used to monitor the reading progress of students is oral reading fluency (ORF). This study examined the extent to which ORF slope predicts reading comprehension outcomes for fifth-grade struggling readers ($n = 102$) participating in an intensive reading intervention. Quantile regression models showed that ORF slope significantly predicted performance on a sentence-level fluency and comprehension assessment, regardless of the students' reading skills, controlling for initial ORF performance. However, ORF slope was differentially predictive of a passage-level comprehension assessment based on students' reading skills when controlling for initial ORF status. Results showed that ORF explained unique variance for struggling readers whose post-test performance was at the upper quantiles at the end of the reading intervention but slope was not a significant predictor of passage-level comprehension for students whose reading problems were the most difficult to remediate.

Multitiered frameworks for instruction and assessment, such as response to intervention, are widely implemented in school settings in the primary grades and beyond (Vaughn & Fletcher, 2010) as a means of providing timely prevention and intervention for students at risk for learning disabilities (e.g., Fuchs & Fuchs, 2006; Vaughn, Linan-Thompson, & Hickman, 2003). Teachers implementing these models typically screen all students at the beginning of the year to identify students who require intervention and collect ongoing

progress-monitoring data for students struggling to perform at the benchmark level. Within multitiered instructional systems, progress monitoring is a key mechanism for determining whether a student demonstrates an adequate response to instruction and for evaluating the effectiveness of instruction (Christ, 2008; Fuchs, 2004; Fuchs & Fuchs, 2006; Stecker, Fuchs, & Fuchs, 2005; Vaughn & Fuchs, 2003). Research findings show that systematic progress monitoring leads to improved identification of struggling students, better-designed instructional programs, and greater student achievement (Fuchs & Fuchs, 2002; Fuchs, Fuchs, Hamlett, & Phillips, 1994; Stecker et al., 2005). Thus, progress monitoring is critical for students at risk for reading disabilities, given the substantial evidence that reading difficulties can be prevented through early intervention (e.g., Torgesen, 2004).

Despite evidence that monitoring student progress is associated with improved student performance, questions remain about the most appropriate method (e.g., intercept or slope) for evaluating response to instruction. A basic assumption of multitiered delivery models is that frequent student assessments provide teachers with useful data to determine whether students adequately respond to instruction. Many well-regarded institutions that provide guidance to educators suggest that educators use growth data to inform decisions about students' responsiveness. For instance, the National Center on Student Progress Monitoring states, "Progress toward meeting the student's goals is measured by comparing expected and actual rates of learning" (www.studentprogress.org/progresmon.asp#2). This definition implies that slope, defined as the trend of a student's performance over time on general outcome measures or mastery assessments, is the most appropriate criteria for evaluating response to instruction; however, the research on the predictive value of slope data is equivocal (e.g., Schatschneider, Wagner, & Crawford, 2008; Tolar, Barth, Fletcher, Francis, & Vaughn, 2014).

The purpose of this study was to better understand the value of using slope data to monitor student progress by examining the predictive validity of oral reading fluency (ORF) on reading comprehension of struggling fifth-grade readers. We describe ORF measures and their relation to reading comprehension, review the past research on the predictive utility of ORF slope data for reading comprehension outcomes, explain how our study enhances understanding of monitoring student progress, and pose two research questions.

What Does ORF Measure?

ORF and reading comprehension.

ORF is a widely used measure for screening students and monitoring student progress in reading (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003). During ORF assessments, students read connected text for a given amount of time, and the number of words read correctly per minute is calculated as an index of reading fluency. Although ORF directly measures students' reading fluency of connected text, researchers also suggest that reading fluency is an indicator of overall reading competence (Fuchs et al., 2001; Jenkins et al., 2003). Reading theories and empirical findings help to substantiate the relation between ORF and reading comprehension. The automaticity model of reading (LaBerge & Samuels, 1974), the interactive-compensatory model of reading (Stanovich, 1980), and the verbal efficiency model of reading (Perfetti, 1985) share the view

that the student's efficiency of translating written words to spoken language facilitates reading comprehension. These theories suggest that efficient sublexical and lexical processes free up attentional capacity for complex comprehension processes. For instance, a student who instantly recognizes text has increased attentional capacity for complex comprehension processes, such as relating new information with existing background knowledge or filling knowledge gaps through inference-making. There is also consistent empirical support for the strong relation between ORF and reading comprehension, especially from kindergarten through grade 3. One meta-analysis (Reschley et al., 2009) found strong correlations (.6 to .7) between ORF and reading achievement for students in grades 1–6, with lower correlations for students in grades 5 and 6, although the grade-level differences were not statistically significant.

ORF beyond primary grades.

Prior research suggests that the relation between ORF and reading comprehension may vary based on students' developmental phase of reading (e.g., Yovanoff, Duesbery, Alonzo, & Tindal, 2005). ORF is a particularly good indicator of reading comprehension during the beginning reading stage (Fuchs et al., 2001). During the initial stages of learning to read, word-level skills are central to students' overall reading competence. However, in the upper-elementary grades, language abilities play a greater role than word-level reading skills in reading comprehension (Vellutino, Tunmer, Jaccard, & Chen, 2007). Reading, often represented as a unitary construct in the early elementary grades, manifests as two distinct factors (word-level reading and language abilities) in the upper-elementary grades (Shinn et al., 1992; Storch & Whitehurst, 2002). When reading is represented as two distinct factors, ORF is often described as a measure of word-level reading skills rather than a measure of reading comprehension (Cirino et al., 2013; Shinn et al., 1992).

Although previous studies have documented strong bivariate correlations between ORF and reading comprehension for students beyond the early elementary grades, findings also suggest a decline in the strength of the association between ORF and reading comprehension as students move beyond grade 3 (Baker et al., 2015; Barth et al., 2012; Espin & Foegen, 1996; Espin, Wallace, Lembke, Campbell, & Long, 2010; Ticha, Espin, & Wayman, 2009; Yovanoff et al., 2005). Studies show that measures of language skills, such as vocabulary and maze tasks, are more predictive than ORF for comprehension in older students (Denton et al., 2011; Espin & Foegen, 1996; Shapiro, Solari, & Petscher, 2008; Stevenson, 2015; Yovanoff et al., 2005). For these reasons, there is no consensus on whether ORF is an effective progress-monitoring measure for predicting reading comprehension of students in the upper-elementary grades (Denton et al., 2011; Fuchs, Fuchs, & Maxwell, 1988; Graney, Martinez, Missall, & Aricak, 2010; Jenkins & Jewell, 1993; Wayman, Wallace, Wiley, Ticha, & Espin, 2007), and thus, further study is warranted.

Predictive Utility of ORF Slope

There are several methods for assessing students' response to instruction (e.g., Fuchs & Deshler, 2007; Torgesen, 2000; Vellutino et al., 1996). Fuchs and Fuchs (1998) put forth three methods for evaluating response to instruction: (1) final status, (2) slope discrepancy,

and (3) dual discrepancy. Final status compares the student's final score on a norm-referenced measure to an established cut-point, such as the 20th percentile. Using slope discrepancy, teachers compare the rate of growth for an individual student with the normative rate of growth. The dual discrepancy method combines final status and slope to assess a student's instructional response. Selecting an appropriate method for assessing a student's responsiveness to instruction requires identifying the method with the highest predictive utility and considering the resource demands of each method. For instance, if slope discrepancy provides significantly more predictive value than final status, it may be advisable for teachers to dedicate additional resources to collecting multiple progress-monitoring assessments and evaluating growth over time.

We identified only six studies that have examined the predictive utility of ORF slope on reading comprehension performance (Baker et al., 2008; Kim, Petscher, Schatschneider, & Foorman, 2010; Schatschneider et al., 2008; Tolar et al., 2014; Wanzek et al., 2010; Yeo, Fearing, & Christ, 2012). Four studies evaluated the predictive utility of ORF slope for reading outcomes among students in grades 1–3. Baker et al. examined how initial intercept and slope of ORF estimated across grades predict end-of-grade state reading achievement test, in two cohorts of students, from grades 1 to 2 (Cohort 1) and grades 2 to 3 (Cohort 2). They found that both ORF intercept and slope predicted performance on the state assessments from grades 1 to 2 and grades 2 to 3. However, only 3% of the additional variance was explained by slope from grades 2 to 3, compared to 10% unique variance explained in Cohort 1. Wanzek et al. found that slope estimated using data collected in grades 1–3 predicted state reading test performance in grade 3 when controlling for initial ORF status. Similarly, Kim et al. reported that ORF slope significantly predicted end-of-year reading achievement beyond initial benchmark status (Fall ORF score) among grade 1 students. However, ORF slope was not predictive for the same cohort of students in grades 2–3 beyond initial or final benchmark status. Further, Schatschneider et al. found that ORF linear slope was not predictive of reading comprehension for grade 1 students when controlling for final ORF status. These studies suggest that ORF slope in the early grades may have predictive power for reading comprehension beyond initial status but not beyond final status.

Two studies examined the slope validity for students in the later grades. Yeo et al. (2012) analyzed whether ORF slope predicted scores on an end-of-year state reading achievement test for students in grades 3–8. They found that ORF slope did not predict reading achievement for any of these grade levels when accounting for initial ORF status. Tolar et al. (2014) found that ORF slope was a significant predictor of reading comprehension beyond initial status for typical students in grades 6–8, as identified by a state reading achievement test. Interestingly, ORF slope did not add to the prediction of reading comprehension beyond the initial ORF status for the subsample of struggling readers, for whom progress monitoring is most needed. Tolar et al. also examined the alignment between a progress-monitoring measure and its targeted outcome, under the assumption that greater alignment would yield greater predictive utility. Findings showed that ORF slope was predictive for both typical and struggling readers when the reading outcome (i.e., sight word reading measure) highly aligned with the progress monitoring measure. These findings suggest that the predictive utility of ORF slope may vary based on (a) student grade level, (b) whether initial or final

progress-monitoring status is modeled, (c) the reading comprehension ability of students (i.e., typical or struggling readers), and (d) the alignment of the progress-monitoring measure and the outcome.

Current Study

This study is situated within a larger randomized control trial that examined the effects of 2 years of intensive intervention for struggling readers in grades 4 and 5 (Vaughn, Solis, Miciak, Taylor, & Fletcher, 2016). We used a subsample of students assigned to 2 years of treatment and collected progress-monitoring data in the second year of treatment. The primary purpose of this study was to examine the predictive utility of ORF slope on reading comprehension beyond the information provided by ORF initial performance. We reasoned that if the most intensive level of instruction in multitiered approaches to instruction and assessment call for the consideration of growth when evaluating students' responsiveness, ORF slope should demonstrate unique predictive value for students who are most at risk for reading difficulties.

We explored the possibility of differential predictive validity of slope based on students' reading level and the degree of alignment between progress monitoring measure and criterion assessment. Past research suggests that ORF slope was differentially predictive based on students' reading level. We extend past research by using quantile regression to better understand the association between ORF slope and reading comprehension as a function of students' responsiveness (as indicated by their comprehension posttest scores rather than creating separate groups using an arbitrary cut-point). This method can shed light on how the magnitude of the predictive relation between ORF growth and reading comprehension may vary at several points along the distribution of end-of-year comprehension scores. In addition, based on Tolar et al.'s (2014) finding that the predictive utility of slope is highest when the criterion measure aligns with the progress monitoring measure, we evaluated the effect of ORF slope on a sentence-level fluency and comprehension assessment (high alignment) and passage-level reading comprehension assessment (low alignment). The primary purpose of this study was to explore possibilities differential prediction of ORF slope based on the reading level and the level of alignment between criterion measures. To address the following research questions we used the individual's best fitting line estimated from individual growth curve analysis to reflect the current school practices of calculating slope (e.g., NCS Pearson, 2012):

1. Does ORF growth predict reading outcomes for struggling fifth-grade readers beyond the information provided by initial performance?
2. How do the predictive relations between ORF growth and reading outcomes vary based on the alignment of the progress-monitoring measure and outcome and the reading ability of students?

Method

Participants

Struggling readers were initially screened using a cut-point at or below a standard score of 85 on the Gates-MacGinitie Reading Test (GMRT; Gates, 2000) in the fall of the previous year (when they were in fourth grade). 148 students were randomly assigned to 2-year treatment. ORF data were collected for 117 students (at least one data point). Among these students, 15 students' GMRT posttest scores were unavailable. An examination of missing data patterns indicated that the data were missing completely at random (Littles MCAR test = 6.232, $df=8$, $p=.621$). Multiple imputation was conducted in Stata 13.1 (StataCorp, 2013) with a multivariate normal model using measures of verbal IQ, sight word efficiency, pretest GMRT creating 100 imputed datasets to be averaged. Analyses with multiple imputation yielded similar patterns of results; therefore, we present results from the original dataset of 102 students. The demographic information of participants is summarized in Table 1. In this sample, approximately half of the students were female (51.96%) and identified as being limited English proficient (55.88%), and a majority of the students received free or reduced-price lunch (90.20%).

Measures

Reading comprehension.—The Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010) was administered as a measure of reading fluency and comprehension. The TOSREC is a 3-minute, group-administered test that requires students to silently read and evaluate the truthfulness of a series of sentences. The TOSREC has different forms to be administered at different times in the academic year. We administered Form A in the spring, when Form C is typically administered; therefore, we report raw scores in our analyses. The TOSREC is a measure of both reading fluency and comprehension. Findings show that the TOSREC is more strongly related to reading comprehension than other measures of silent reading fluency (Denton et al., 2011) and measures of ORF (Johnson, Pool, & Carter, 2011). Yet as a sentence-level measure of fluency and comprehension in which students read simple sentences such as “People have teeth” or “The sky is always brown and yellow,” the TOSREC has been modeled as a distinct construct from reading comprehension that involves higher-order language skills for older students (Cirino et al., 2013).

Students also completed the fourth edition of the GMRT, as part of their initial screener. Similar to many statewide reading achievement tests, the GMRT is a group-administered test of reading in which students read expository and narrative passages of 3 to 15 sentences and answer 48 multiple-choice comprehension questions. Previous research examining grade 5 students shows performance on the GMRT is more highly correlated with performance on tests that assess components of higher-level reading comprehension (e.g., inferencing) than on measures of fluency and comprehension like the TOSREC (Klauda & Guthrie, 2008).

AIMSweb.—The AIMSweb curriculum-based measure of oral reading (Shinn & Shinn, 2002) assesses ORF for grade-level connected texts. Students read a passage aloud for 1 minute, and the number of words read accurately is calculated. Hesitations of more than 3

seconds and words omitted or substituted are scored as errors. Words self-corrected within 3 seconds are regarded as correct. In this study, the same passage was administered at each progressmonitoring time point. We decided to use the same ORF form to monitor student progress, rather than alternate forms. This decision was based on research that shows that different forms lead to inaccurate estimates of student growth due to difficulty differences among alternate forms (Francis et al., 2008). Research also shows that a single passage is as predictive of reading comprehension for older students as the mean or median from three passages (Baker et al., 2015); therefore, we administered a single passage at each progress-monitoring time point. Due to relatively small growth in ORF for older students (e.g., Vaughn & Fletcher, 2010), we assessed students on a bimonthly schedule. An investigation of the progress-monitoring practices found that the majority of schools under study monitored progress quarterly or less frequently (Mellard, McKnight, & Woods, 2009). In addition, Jenkins and Terjeson (2011) compared three progress monitoring schedules (every 2, 4, or 8 weeks) and concluded that slope calculated with ordinary least squares yielded similar results regardless of the monitoring schedule. Thus although recommendations suggest more frequent progress monitoring (Jenkins, Graff, & Migliorett, 2009; Safer & Fleischman, 2005), a bimonthly schedule aligns well with current educational practices and enables this study to inform practitioners.

Intervention

Research staff members delivered a standardized intervention using scripted protocols for 30 minutes each day for approximately 16 weeks (November to April). Tutors delivered multicomponent reading instruction targeting vocabulary, word reading, fluency, and reading comprehension using science texts. To improve foundational reading skills, students received phonics instruction with ample opportunities for teacher feedback. Additionally, in text-based reading activities, students summarized read texts and synthesized information across texts. Since the focus of this study is on the use of progress monitoring measures, the reader is referred to another study for additional information about the intervention (Vaughn et al., in press).

Analyses

Ordinary least square vs. quantile regression.—Quantile regression (Koenker & Bassett, 1978) is an alternative analytic technique to ordinary least square (OLS) regression. OLS regression estimates average relations between predictor(s) and an outcome by minimizing the sum of squared residuals. Quantile regression allows researchers to estimate the relations of predictor(s) to an outcome at several points along the continuous distribution of the outcome variable by minimizing a sum of absolute residuals (known as least absolute value). Quantile regression captures a more complete picture by examining how the relation of predictor(s) with an outcome differs at multiple points (or quantiles) in the distribution of the outcome. Quantile regression is particularly useful to study a specific portion of the distribution, such as struggling readers, for statistical and theoretical reasons. First, data from a sample of students with reading difficulties are seldom normally distributed, and quantile regression does not assume homoscedasticity of the error terms and is robust to violations of normality. Second, predictions about students in the lower end of the distribution may be different from predictions about students in other parts of the

distribution. For example, comprehension and ORF are more strongly correlated for more fluent readers in primary grades than for less fluent readers (e.g., Petscher & Kim, 2011). Similarly, quantile regression can help to identify which variables are most predictive of reading comprehension outcomes at various points along the outcome variable. With this information, researchers can suggest specific screening and progress-monitoring measures that are most appropriate for students at specific levels of performance. It is important to note that quantile regression is different from fitting a separate OLS regression for each quantile range. Quantile regression takes advantage of an asymmetric weighting system of data points based on their distance from the specified quantile and uses all data points when estimating the coefficient at a certain percentile.

Analytic steps.—We used a two-step process to address our research questions. As a first step, to obtain estimates of student growth in ORF, we estimated an individual growth curve model of random linear slope and random intercept with ORF scores measured at three time points. Results of the growth model are presented in Table 2. The fixed effect for slope, or the average growth observed per 2-month interval, was 7.25 (i.e., 3.625 words per month). The fixed effect for the initial intercept indicated that students, on average, read 89.45 words correctly per minute at the start of intervention. There was statistically significant variance in intercept and slope across students, with reliability estimates of .91 for the intercept and .33 for the slope. We stored the best linear unbiased prediction of the random effects and used it in subsequent regression models. Once the initial level and slope estimates were saved, we standardized the variables using sample's descriptive statistics prior to including them in prediction models. Then, we ran a series of regression models separately for the TOSREC and GMRT to examine the predictive validity of ORF growth (linear slope). Although this two-step process is not the most rigorous methods of testing predictive validity of slope (as compared to latent growth curve analysis), this method is deemed appropriate because we wanted to reflect the current approach to calculating and using slope practiced by teachers. For example, individual OLS estimates, termed as the best fitting line) is what AIMSweb online graphing tool utilizes.

Following the three methods (i.e., final status, slope discrepancy, and dual discrepancy) for evaluating response to instruction (Fuchs & Fuchs, 1998), we initially examined predictive validity of ORF initial performance (Model 1) and ORF slope (Model 2) separately. Then, we included ORF initial performance and slope simultaneously in the final model (Model 3) to examine whether ORF growth has additional predictive value beyond single time point assessment. The models were initially fit using the OLS method and the models were run using quantile regression analyses with 5000 bootstrap replications of the data to estimate standard error.

Quantile regression model results are presented in four ways. First, we estimated coefficients at three quantiles (.25, .5, and .75). Based on the GMRT posttest standard scores presented in Table 3, the .25 quantile in our sample corresponded approximately to the 16th percentile in the normative population (standard score of 84), the .50 quantile corresponded to the 20th percentile (standard score of 89), and the .75 quantile corresponded to the 30th percentile (standard score of 94). These quantiles are practically meaningful, as the 16th percentile in the normative population is often used as a cut-point for identifying inadequate responders

who may require special education services (e.g., Compton, Fuchs, Fuchs, Lambert, & Hamlett, 2012; Fletcher et al., 2011) and the 25th percentile is frequently applied as the criterion for determining adequate response to intervention (e.g., Torgesen, 2000). Second, we tested for differences in selected quantiles using Wald tests—testing whether the difference of the coefficients in two selected quantiles is zero. Third, we calculated variance in outcome uniquely explained by ORF slope by using the hierarchical regression approach, as suggested by Petscher, Logan, and Zhou (2013). Fourth, we ran simultaneous quantile regressions of Model 3 at increments of .5 quantiles, starting from .15 to .85, and plotted the resulting coefficients to indicate the magnitude of predictive validity of ORF initial status and slope for reading comprehension (Figures 1–4). All models were fit using Stata 13.1 (StataCorp, 2013).

Results

Descriptive Statistics

Descriptive statistics are presented in Table 3. Our sample of struggling readers, initially reading below or at the 16th percentile, were reading at approximately the 5th to the 50th percentiles in GMRT standard scores at the end of 2 years of intervention. In terms of ORF initial performance and slope, we had large variation in our sample. Based on the ORF norms published by Hasbrouck and Tindal (2006), a majority of our sample fell within the 10th to 50th percentile range in terms of initial performance. Our students read 3.58 more words per month, which approximates the average growth rate of fifth-graders (.9 words per week) and slightly exceeds the average growth rate for fifth-grade struggling readers (.7 to .8 words per week) in the normative sample. These data suggest that our struggling readers were generally responsive to intervention in terms of developing foundational reading skills. Except for the ORF slope, which was positively skewed ($p = .006$), all other variables were normally distributed.

Table 4 provides zero-order correlations among the variables. Pearson correlations indicate that the TOSREC and GMRT were only moderately correlated ($r = .26$, $p = .008$). ORF initial performance and slope demonstrated a moderate correlation to the GMRT ($r = .32$, $p < .001$; $r = .23$, $p = .019$, respectively) and a slightly stronger correlation to the TOSREC ($r = .44$, $p < .001$; $r = .37$, $p < .001$, respectively). ORF initial performance and slope were also moderately correlated ($r = .47$, $p = .02$).

OLS and Quantile Regression

Table 5 presents a comparison of standardized regression coefficients obtained from OLS regression and quantile regression analyses. Results of the quantile regression analyses are also depicted in Figures 1–4. In these figures, the x -axes refer to the level of comprehension (presented from the 10th percentile to 90th percentile) and the y -axes represent the regression coefficients. The error bars around the estimates refer to the 90% confidence interval.

TOSREC.—OLS regression of Model 1 and Model 2 indicated that ORF initial performance ($\beta = .38$, $p < .001$) and slope ($\beta = .35$, $p < .001$) were significant predictors of

the TOSREC independently. In Model 3, ORF initial performance ($\beta = .31, p = .001$) and slope ($\beta = .27, p = .003$) were significant predictors, controlling for each other. Results of the quantile regression analyses provided similar patterns. For Model 1, ORF initial performance was a significant predictor of the TOSREC at all three quantiles, although the association between ORF initial performance and the TOSREC tends to get stronger with higher levels of performance on the TOSREC (β s = .32, .41, and .65 at .25, .5, and .75 quantiles). For Model 2, ORF slope was a significant predictor of the TOSREC and coefficients remained constant across the distribution (β s = .36, .24, and .30 at .25, .5, and .75 quantiles). As indicated in Model 3, the associations of the TOSREC with ORF initial performance and slope remained statistically significant after controlling for each other (Figures 1 and 2). Most of the comparisons between quantiles did not indicate statistically significant differences, with one exception. ORF initial performance was more strongly predictive of the TOSREC at the .75 quantile than at the .25 quantile (Table 6). Additionally, 4% to 5% of the variance in the TOSREC was uniquely predicted by ORF slope across the selected quantiles.

GMRT.—OLS regression of Model 1 and Model 2 indicated that ORF initial performance ($\beta = .29, p < .003$) and slope ($\beta = .20, p < .039$) are significant predictors of the GMRT independently. However, when ORF initial performance was included in Model 3, the coefficient for slope did not remain significant ($\beta = .14, p = .149$) controlling for the significant association with ORF initial performance ($\beta = .26, p = .009$). Quantile regression analyses showed the predictive utility of ORF slope was conditional on reading comprehension level. For Model 1, ORF initial performance was a significant predictor of the GMRT at the lower quantiles but not significant at the upper quantile (β s = .44, .33, and .21 at .25, .5, and .75 quantiles). A contrasting pattern was found for Model 2. ORF slope was not a significant predictor of the TOSREC at the lower quantile but was significant at the upper quantiles (β s = .06, .31, and .31 at .25, .5, and .75 quantiles). As depicted in Figures 3 and 4, ORF initial performance was more likely to be statistically significant at the lower quantile of GMRT distribution, whereas ORF slope was more likely to be statistically significant at the upper quantile of the GMRT. However, differences between the selected two quantiles were not statistically significant (Table 6). Variance explained uniquely by ORF slope also increased at upper quantiles. At the .25 quantile, slope explained less than 1% of the variance in the GMRT, but it increased to 2% to 3% at the .5 and .75 quantiles.

Discussion

Prior research has documented ORF as a reliable and valid measure for screening at-risk students with reading difficulties and progress monitoring to determine students' response to instruction in multitiered systems (e.g., Kovaleski, VanDerHeyden, & Shapiro, 2013). We tested the assumption that ORF slope provides additional predictive information beyond a single-time-point assessment for a sample of struggling readers in fifth grade. The present study focused on struggling readers in fifth grade because the goals of reading instruction shift from foundational to higher-order comprehension skills in the upper-elementary grades and many struggling readers fail to meet these changing demands (Chall & Jacobos, 2003). We used quantile regression to explore the conditional predictive validity of ORF slope for

explaining reading comprehension as a function of students' responsiveness to intervention, as indicated by their comprehension posttest scores. Additionally, we explored the effects of alignment between the progress monitoring and the outcome on the predictive validity of ORF slope by using a sentence-level fluency and comprehension measure (high alignment) and a passage-level reading comprehension measure (low alignment).

Predictive Validity of ORF Slope for Struggling Readers

ORF slope was a significant predictor of sentence-level fluency and comprehension when controlling for initial ORF performance, a finding that is consistent with previous research on the alignment of a performance on progress monitoring measure and its targeted outcome (Tolar et al., 2014). Our findings suggest that the predictive validity of ORF slope was relatively constant for students at all reading levels when used to predict the highly aligned measure, suggesting that there are no differences between the predictive validity of ORF based on student performance. However, we found a different pattern of results when evaluating the predictive validity of ORF slope for passage-level reading comprehension. Consistent with previous research on the predictive validity of ORF slope for students in grades 3–8 (Tolar et al., 2014; Yeo et al., 2012), ORF slope did not significantly add to the prediction when the outcome measure was passage-level reading comprehension. However, ORF slope showed predictive validity based on students' end-of-year passage-level reading comprehension performance for students who made relatively strong progress (above .75 quantile).

These findings suggest that ORF slope is predictive of student performance on proximal measures of reading fluency and comprehension for all fifth-grade struggling readers, but is only predictive of passage-level reading comprehension for struggling readers who made relatively strong progress. The results provide evidence that the level of alignment between the progress monitoring measure and outcome affects the predictive validity of progress-monitoring slope. As expected, ORF slope demonstrated greater predictive value when the outcome was highly aligned.

It is not apparent why ORF slope was predictive on passage-level comprehension for students who responded at a higher rate to the intervention, yet was not uniquely predictive for most struggling readers. One possible explanation is that for students in our study who continued to struggle with higher-level reading comprehension tasks, growth in reading fluency may not have transferred to improved passage-level reading comprehension because reading comprehension requires more complex linguistic skills and background knowledge. Although theory suggests that growth in ORF is associated with greater cognitive resources for higher-order cognitive processes (e.g., Perfetti, 1985), it may be that these students lacked other attributes (e.g., background knowledge and inference skills) required to perform well on a passage-level reading comprehension measure (Cromley & Azevedo, 2007; Kintsch, 1994; RAND, 2002). In a study of the cognitive profiles of adequate and inadequate responders with the same sample of students, Cho et al. (2015) observed that language-related areas (verbal knowledge and listening comprehension) were significantly lower in inadequate responders than in adequate responders. Thus, students who responded well to the intervention may have had sufficient language-related skills to benefit from the

additional cognitive resources freed up through improved reading fluency whereas students in the lower quartiles did not have such resources.

It is also worth noting that both intercept and slope of ORF were only moderately correlated with reading comprehension ($.23 < rs < .44$) and intercept and slope together explained only 10 – 14% of the variance in sentence-level reading fluency and comprehension across quartiles and 5 - 9% of the variance in passage-level reading comprehension. Amount of variance slope uniquely explained were also limit to 2 – 5% across the models. The small *R*-squares across the models suggest that ORF does not explain much of the variance in reading comprehension outcomes for these upper elementary struggling readers. This finding is in line with prior research demonstrating that a combination of fluency, vocabulary, and comprehension measures leads to higher predictive accuracy of reading comprehension problems for students in upper grades than solely relying on ORF (e.g., Baker et al., 2015; Nese, Park, Alonzo, & Tindal, 2011; Yavanoof et al., 2005).

Limitations and Future Research

We note several limitations when interpreting the results. First, although our sample size ($n = 102$) is comparable to previous quantile regression studies (e.g., Tighe & Schatschneider, 2014), it is nonetheless small, which affects the statistical power and stability of our model and results. Standard errors of the coefficients tend to be large especially at the extreme quantiles (i.e., $> .70$ quantile or $< .30$ quantile) in quantile regression models (Petscher et al., 2013). Our small sample size may have exacerbated this problem, resulting in statistically nonsignificant differences between quantile comparisons. In addition, the small sample size limited the number of predictors included in the model. Second, pretest performance was not included in the model due to the constraints in our research design as well as the sample size limitation. We used GMRT pretest as a screener for the intervention study; thus, the distribution of GRMT pretest was systematically restricted, which could cause ORF (unrestricted distribution) to explain greater variance in posttest GMRT performance than pretest GMRT performance. Additionally, to make the models for each outcome comparable, we did not include TOSREC pretest in the TOSREC model. Third, ORF assessments were administered less frequently than is typical when students receive an intensive intervention in a multitiered instructional system. More frequent progress monitoring—such as twice a month, as is recommended for students receiving an intensive intervention (Jenkins & Terjeson, 2011)—would have better aligned with current recommendations and may have yielded a more reliable estimate of slope. Fourth, the reliability of the slope estimate was low (.39). Although this slope estimate is consistent with the previous study (Tolar et al., 2014), a more reliable slope estimate may have increased the power to detect significant slope effects on outcomes. Fifth, it is noteworthy that the goal of this study was to explore differential predictive validity of ORF slope and to find empirical support of the current progress monitoring practice for upper elementary grades in schools, and not to build the most parsimonious model to test the theoretical model. Although latent growth modeling allows us to model growth estimates without measurement error and use them to predict outcome in a single model, the models presented in this study more accurately aligns with current school practices.

Our study explored the possibility of differential predictive validity of ORF slope with limited sample size. We note several avenues for future research. First, our results indicate overlapping confidence intervals across the quantiles and statistically nonsignificant differences on many of the between quantile comparisons. The statistically nonsignificant coefficients when directly compared could be in part due to our small sample size as well as restriction of range as we only included struggling readers. Thus, although this may indicate an appropriateness of non-quantile models, a larger sample that includes heterogeneous reading levels might detect statistically significant differences across the quantiles. Second, future research should investigate whether the predictive validity of progress-monitoring slope differs as a function of the frequency of progress monitoring or progress-monitoring measures. Finally, future research on the predictive validity of progress-monitoring slope should consider other progress-monitoring measures such as vocabulary or maze tasks, which assess comprehension-related skills more directly.

Implications for Practice

With these caveats in mind, findings from the present study suggest that if the progress-monitoring measure is highly aligned with the outcome of interest, using slope information may lead to better understanding of students' responsiveness to instruction than just using initial performance. Our results do not provide evidence that ORF slope data are useful when predicting response among fifth-graders with persistent reading difficulties on outcomes of higher-level reading comprehension skills. These results highlight the importance of carefully selecting a progress-monitoring measure based on the reading outcome of interest. Past research suggests that a student's most recent performance on a progress-monitoring measure is the most reliable predictor of performance on a standardized reading achievement test (Schatschneider et al., 2008). This study's findings do not test this notion; therefore, practitioners may prefer to use the most recent progress-monitoring score, rather than calculating and using the rate of student growth, when making predictions about a student's future performance.

Some researchers have questioned the validity of ORF as a predictor of higher-level reading comprehension (e.g., Parker, Hasbrouck, & Tindal, 1992). For instance, Fuchs et al. (2001) suggest that ORF is actually an indicator of "basic reading competence." Our findings echo this concern. Although we did not include other progress-monitoring measures to compare with ORF, prior research suggests that maze tests may be more appropriate for monitoring the reading comprehension progress of older students. For example, Espin et al. (2010) and Ticha et al. (2009) found that ORF was not associated with change in reading comprehension performance for older students, whereas growth on maze tasks was related to performance on reading achievement tests and changes on standardized readings. Thus, practitioners may consider using various progress-monitoring measures (e.g., ORF and maze tasks) aligned to both basic reading skills and higher-order comprehension skills for upper-elementary students receiving intensive interventions.

Acknowledgments

This research was supported by grant P50 HD052117-07 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not

necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health.

References

- Baker DL, Biancarosa G, Park BJ, Bousselot T, Smith J-L, Baker SK, ... Tindal G (2015). Validity of CBM measures of oral reading fluency and reading comprehension on high-stakes reading assessments in grades 7 d 8. *Reading and Writing*, 28(1), 57–104. doi: 10.1007/s11145-014-9505-4
- Baker SK, Smolkowski K, Katz R, Fien H, Seeley JR, Kame'Enui EJ, & Beck CT (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review*, 37, 18–37.
- Barth AE, Stuebing KK, Fletcher JM, Cirino PT, Romain M, Francis D, & Vaughn S (2012). Reliability and validity of oral reading fluency median and mean scores among middle grade readers when using equated texts. *Reading Psychology*, 33(1–2), 133–161. doi: 10.1080/02702711.2012.631863 [PubMed: 23087532]
- Chall JS, & Jacobs VA (2003). The classic study on poor children's fourth-grade slump. *American educator*, 27(1), 14–15.
- Christ TJ (2008). Best practices in problem analysis In Thomas A & Grimes J (Eds.), *Best practices in school psychology* (pp. 159–176). Bethesda, MD: National Association of School Psychologists.
- Cho E, Roberts GJ, Capin P, Roberts G, Miciak J, & Vaughn S (2015). Cognitive Attributes, Attention, and Self-Efficacy of Adequate and Inadequate Responders in a Fourth Grade Reading Intervention. *Learning Disabilities Research & Practice*, 30, 159–170. [PubMed: 26997755]
- Cirino PT, Romain MA, Barth AE, Tolar TD, Fletcher JM, & Vaughn S (2013). Reading skill components and impairments in middle school struggling readers. *Reading and Writing*, 26(7), 1059–1086. doi:10.1007/s11145-012-9406-3 [PubMed: 24000271]
- Compton DL, Fuchs LS, Fuchs D, Lambert W, & Hamlett C (2012). The cognitive and academic profiles of reading and mathematics learning disabilities. *Journal of Learning Disabilities*, 45(1), 79–95. doi:10.1177/0022219410393012 [PubMed: 21444929]
- Cromley JG, & Azevedo R (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2), 311–325. doi: 10.1037/0022-0663.99.2.311
- Denton CA, Barth AE, Fletcher JM, Wexler J, Vaughn S, Cirino PT, ... Francis DJ (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading*, 15(2), 109–135. doi:10.1080/10888431003623546 [PubMed: 21637727]
- Espin CA, & Foegen A (1996). Validity of general outcome measures for predicting secondary students' performance on content-area tasks. *Exceptional Children*, 62(6), 497–514.
- Espin C, Wallace T, Lembke E, Campbell H, & Long JD (2010). Creating a progress-monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research & Practice*, 25(2), 60–75. doi: 10.1111/j.1540-5826.2010.00304.x
- Fletcher JM, Stuebing KK, Barth AE, Denton CA, Cirino PT, Francis DJ, & Vaughn S (2011). Cognitive correlates of inadequate response to reading intervention. *School Psychology Review*, 40(1), 3–22. [PubMed: 23125475]
- Francis DJ, Santi KL, Barr C, Fletcher JM, Varisco A, & Foorman BR (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology*, 46(3), 315–342. doi:10.1016/j.jsp.2007.06.003 [PubMed: 19083362]
- Fuchs D, & Deshler DD (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice*, 22(2), 129–136.
- Fuchs LS, & Fuchs D (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice*, 13(4), 204–219.
- Fuchs LS, & Fuchs D (2002). Curriculum-Based Measurement: Describing Competence, Enhancing Outcomes, Evaluating Treatment Effects, and Identifying Treatment Nonresponders. *Peabody Journal of Education*, 77, 64–84.

- Fuchs D, & Fuchs LS (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41(1), 93–99. doi:10.1598/RRQ.41.1.4
- Fuchs LS (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, 33(2), 188–193.
- Fuchs LS, Fuchs D, Hamlett CL, & Phillips NB (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children*, 60, 518–537.
- Fuchs LS, Fuchs D, Hosp MK, & Jenkins JR (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5(3), 239–256. doi:10.1207/S1532799XSSR0503_3
- Fuchs LS, Fuchs D, & Maxwell L (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, 9(2), 20–28. doi:10.1177/074193258800900206
- MacGinitie WH (2000). *Gates-MacGinitie reading tests*. Itasca, IL: Riverside.
- Graney SB, Martinez RS, Missall KN, & Aricak OT (2010). Universal screening of reading in late elementary school R-CBM versus CBM maze. *Remedial and Special Education*, 31(5), 368–377. doi:10.1177/0741932509338371
- Jenkins JR, Fuchs LS, van den Broek P, Espin C, & Deno SL (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95(4), 719–729. doi:10.1037/0022-0663.95.4.719
- Jenkins JR, Graff JJ, & Miglioretti DL (2009). Estimating reading growth using intermittent CBM progress monitoring. *Exceptional Children*, 75, 151–163.
- Jenkins JR, & Jewell M (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children*, 59(5), 421–432.
- Jenkins J, & Terjeson KJ (2011). Monitoring reading growth: Goal setting, measurement frequency, and methods of evaluation. *Learning Disabilities Research & Practice*, 26(1), 28–35. doi:10.1111/j.1540-5826.2010.00322.x
- Johnson ES, Pool JL, & Carter DR (2011). Validity evidence for the Test of Silent Reading Efficiency and Comprehension (TOSREC). *Assessment for Effective Intervention*, 37(1), 50–57.
- Kim Y-S, Petscher Y, Schatschneider C, & Foorman B (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology*, 102(3), 652–667. doi:10.1037/a0019643
- Kintsch W (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294–303. doi:10.1037/0003-066X.49.4.294 [PubMed: 8203801]
- Klauda SL, & Guthrie JT (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, 100(2), 310–321.
- Koenker R, & Bassett G (1978). Regression quantiles. *Econometrica*, 46(1), 33–50.
- Kovaleski JF, VanDerHeyden AM, & Shapiro ES (2013). *The RTI approach to evaluating learning disabilities*. New York, NY: Guilford Press.
- LaBerge D, & Samuels SJ (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293–323.
- Mellard D, McKnight M, & Woods K (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice*, 24(4), 186–195.
- NCS Pearson (2012). *Aimsweb Introductory Guide*. Bloomington, MN: Pearson Education Inc.
- Parker R, Hasbrouck JE, & Tindal G (1992). The maze as a classroom-based reading measure: Construction methods, reliability, and validity. *The Journal of Special Education*, 26(2), 195–218.
- Perfetti CA (1985). *Reading ability*. New York, NY: Oxford University Press.
- Petscher Y, & Kim Y-S (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *Journal of School Psychology*, 49(1), 107–129. doi:10.1016/j.jsp.2010.09.004 [PubMed: 21215838]
- Petscher Y, Logan JAR & Zhou C (2013). Extending conditional means modeling: An introduction to quantile regression In *Applied Quantitative Analysis in Education and the Social Sciences* (Eds.). New York, NY.

- Reschly AL, Busch TW, Betts J, Deno SL, & Long JD (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47*, 427–469. [PubMed: 19808123]
- Safer N, & Fleischman S (2005). Research matters: How student progress monitoring improves instruction. *Educational Leadership, 62*(5), 81–83.
- Schatschneider C, Wagner RK, & Crawford EC (2008). The importance of measuring growth in response to intervention models: Testing a core assumption. *Learning and Individual Differences, 18*(3), 308–315. doi:10.1016/j.lindif.2008.04.005 [PubMed: 22224065]
- Shapiro ES, Solari E, & Petscher Y (2008). Use of a measure of reading comprehension to enhance prediction on the state high stakes assessment. *Learning and Individual Differences, 18*(3), 316–328. [PubMed: 26347390]
- Shinn MR, Good III RH, Knutson N, Tilly III WD, & Collins VIL (1992). Oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 459–479.
- Shinn MR, & Shinn MM (2002). AIMSweb training workbook: Administration and scoring of reading maze for use in general outcome measurement. Eden Prairie, MN: Edformation.
- Stevenson NA (2015). Predicting Proficiency on Statewide Assessments: A Comparison of Curriculum-Based Measures in Middle School. *The Journal of Educational Research, 108*(6), 492–503.
- Snow CE, Burns MS, & Griffin P (Eds.). (1998). Preventing reading difficulties in young children. Washington, DC: National Academy Press.
- Stanovich KE (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32–71.
- StataCorp LP (2014). Stata 13.0. College Station, Texas: StataCorp LP.
- Stecker PM, Fuchs LS, & Fuchs D (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools, 42*(8), 795–819. doi:10.1002/pits.20113
- Storch SA, & Whitehurst GJ (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934–947. doi: 10.1037//0012-1649.38.6.934 [PubMed: 12428705]
- Tichá R, Espin CA, & Wayman MM (2009). Reading progress monitoring for secondary-school students: Reliability, validity, and sensitivity to growth of reading-aloud and maze-selection measures. *Learning Disabilities Research & Practice, 24*(3), 132–142.
- Tighe EL, & Schatschneider C (2014). A quantile regression approach to understanding the relations among morphological awareness, vocabulary, and reading comprehension in adult basic education students. *Journal of Learning Disabilities*. Advance online publication. doi: 10.177/0022219414556771.
- Torgesen JK (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice, 15*, 55–64.
- Torgesen JK (2002). The prevention of reading difficulties. *Journal of School Psychology, 40*, 7–26.
- Torgesen JK (2004). Avoiding the devastating downward spiral: The evidence that early intervention prevents reading failure. *American Educator, 28*, 6–19.
- Tolar TD, Barth AE, Fletcher JM, Francis DJ, & Vaughn S (2014). Predicting reading outcomes with progress monitoring slopes among middle grade students. *Learning and Individual Differences, 30*, 46–57. doi: 10.1016/j.lindif.2013.11.001 [PubMed: 24659899]
- Vaughn S, & Fletcher JM (2010). Thoughts on rethinking response to intervention with secondary students. *School Psychology Review, 39*, 296–299. [PubMed: 21738283]
- Vaughn S, & Fuchs LS (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*, 137–146.
- Vaughn S, Linan-Thompson S, & Hickman P (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children, 69*(4), 391–409.
- Vaughn S, Solis M, Miciak J, Taylor WP, & Fletcher JM (2016). Effects From a Randomized Control Trial Comparing Researcher and School-Implemented Treatments With Fourth Graders With Significant Reading Difficulties. *Journal of research on educational effectiveness, 9*, 23–44. [PubMed: 28491206]

- Vellutino FR, Scanlon DM, Sipay ER, Small SG, Pratt A, Chen R, & Denckla MB (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88, 601–638.
- Vellutino FR, Tunmer WE, Jaccard JJ, & Chen R (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*, 11(1), 3–32. doi:10.1080/10888430709336632
- Wagner RK, Torgesen JK, Rashotte CA, & Pearson NA (2010). Test of silent reading efficiency and comprehension. Pro-Ed.
- Wanzek J, Roberts G, Linan-Thompson S, Vaughn S, Woodruff AL, & Murray CS (2010). Differences in the relationship of oral reading fluency and high-stakes measures of reading comprehension. *Assessment for Effective Intervention*, 35(2), 67–77. doi:10.1177/1534508409339917 [PubMed: 21479152]
- Wayman MM, Wallace T, Wiley HI, Ticha R, & Espin CA (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, 41(2), 85–120.
- Yeo S, Fearing JY, & Christ TJ (2012). Relation between CBM-R and CBM-mR slopes: An application of latent growth modeling. *Assessment for Effective Intervention*, 37(3), 147–158. doi:10.1177/1534508411420129
- Yovanoff P, Duesbery L, Alonzo J, & Tindal G (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement: Issues and Practice*, 24(3), 4–12.

Figure 1.

Model 3 quantile regression plot of the regression coefficients of ORF initial performance

Note. Error bars indicate 90% confidence intervals.

Figure 2.
Model 3 quantile regression plot of the regression coefficients of ORF slope
Note. Error bars indicate 90% confidence intervals.

Figure 3.

Model 3 quantile regression plot of the regression coefficients of ORF initial performance

Note. Error bars indicate 90% confidence intervals.

Figure 4.
Model 3 quantile regression plot of the regression coefficients of ORF slope
Note. Error bars indicate 90% confidence intervals.

Table 1

Demographics of the Participants

Variable	Count	Percent
Gender		
Female	53	51.96
Male	49	48.04
Race		
African American	17	16.67
Hispanic	33	32.35
White	11	10.78
Other or two or more races	41	40.2
Free or reduced-price lunch		
No	7	6.86
Yes	92	90.20
Missing	3	2.94
Special education		
No	50	29.02
Yes	9	8.82
Missing	43	42.16
Limited English proficiency		
No	45	44.12
Yes	57	55.88

Table 2

Individual Growth Curve Model

	Coefficient	SE	<i>t</i>	<i>p</i>
Fixed effects				
Intercept	89.446	2.398	37.300	< .001
Time	7.253	0.718	10.095	< .001
	Standard deviation	Variance component	Chi-square	<i>p</i>
Random effects				
Intercept	24.885	619.243	1,172.426	< .001
Time	4.168	17.370	148.290	0.001
Residual	8.211	67.416		
Reliability				
Intercept	0.911			
Time	0.329			

Table 3

Descriptive Statistics

	<i>N</i>	<i>M</i>	<i>SD</i>	Minimum	Maximum
GMSS 1	102	77.44	5.63	65	85
GMSS 18	102	87.67	6.89	72	103
TOSREC 1	102	10.92	5.45	0	27
TOSREC 18	102	15.75	6.57	0	31
ORF 14	101	89.46	26.30	20	163
ORF 16	93	97.19	27.83	29	163
ORF 18	90	104.54	28.96	32	180
ORF SL	102	3.58	1.22	1.39	7.83
ORF LV	102	88.96	24.34	25.02	154.80

Note. GMSS 1 = Gates-MacGinitie initial standard score; GMSS 18 = GM posttest standard score; TOSREC 1 = Test of Silent Reading Efficiency and Comprehension initial raw score; TOSREC 18 = TOSREC posttest raw score; ORF 14 = initial oral reading fluency score; ORF 16 = second ORF score; ORF 18 = final ORF score; ORF SL = ORF slope; ORF LV = ORF initial intercept.

Table 4

Zero-Order Correlation

		1	2	3	4	5	6
1	GMSS 18						
2	TOSREC 18	0.26					
3	ORF 14	0.30	0.37				
4	ORF 16	0.30	0.50	0.89			
5	ORF 18	0.34	0.46	0.87	0.91		
6	ORF LV	0.32	0.44	0.97	0.97	0.94	
7	ORF SL	0.23	0.37	0.29	0.54	0.72	0.47

Note. All correlation significant at .05 level. GMSS 18 = Gates-MacGinitie posttest standard score; TOSREC 18 = Test of Silent Reading Efficiency and Comprehension posttest raw score; ORF 14 = initial oral reading fluency score; ORF 16 = second ORF score; ORF 18 = final ORF score; ORF LV = ORF initial intercept; ORF SL = ORF slope.

Table 5

Results of OLS and Quantile Regression Models

Variable	TOSREC					GMRT				
	β	SE	t	P	Pseudo R ²	β	SE	t	P	Pseudo R ²
Model 1										
OLS					.14					.09
Quantile										
25%tile	0.32	0.12	2.77	.01	.07	0.44	0.14	3.20	.00	.08
50%tile	0.41	0.15	2.76	.01	.05	0.33	0.11	3.07	.00	.07
75%tile	0.64	0.17	3.71	.00	.10	0.21	0.19	1.06	.29	.02
Model 2										
OLS					.13					.04
Quantile										
25%tile	0.36	0.12	3.16	.00	.07	0.06	0.09	3.58	.71	.00
50%tile	0.24	0.10	2.5	.01	.05	0.31	0.09	3.27	.00	.06
75%tile	0.30	0.12	2.47	.02	.04	0.31	0.09	2.05	.04	.03
Model 3										
OLS					.21					.11
Quantile										
25%tile					.11					.08
50%tile					.10					.09
75%tile					.14					.05

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Variable	TOSREC					GMRT				
	β	SE	t	P	Pseudo R ²	β	SE	t	P	Pseudo R ²
ORF slope	0.29	0.14	2.13	.04	.04	0.36	0.16	2.30	.02	.03

Table 6

Between Quantile Comparison

		<u>Slope</u>		<u>Level</u>	
Model	Quantile Comparison	$F(1, 98)$	p	$F(1, 98)$	p
TOSREC					
	25 vs. 50	0.04	.85	1.49	.22
	50 vs. 75	0.06	.81	1.47	.23
	25 vs. 75	0.00	.95	4.10	.04
GMRT					
	25 vs. 50	0.52	.47	1.61	.21
	50 vs. 75	2.67	.11	0.21	.65
	25 vs. 75	3.64	.06	1.58	.21