Journal of Educational Psychology
2010, Vol. 102, No. 3, 652–667

0022-0663/10/$12.00 DOI: 10.1037/a0019643

Does Growth Rate in Oral Reading Fluency Matter in Predicting Reading Comprehension Achievement?

Young-Suk Kim
Florida State University and Florida Center for Reading Research, Tallahassee, Florida

Yaacov Petscher
Florida Center for Reading Research, Tallahassee, Florida

Christopher Schatschneider and Barbara Foorman
Florida State University and Florida Center for Reading Research, Tallahassee, Florida

In this study, we examined the relationship of growth trajectories of oral reading fluency, vocabulary, phonological awareness, letter-naming fluency, and nonsense word reading fluency from 1st grade to 3rd grade with reading comprehension in 1st, 2nd, and 3rd grades. Data from 12,536 children who were followed from kindergarten to 3rd grade longitudinally were used. These children were administered Dynamic Indicators of Basic Early Literacy Skills subtests, Peabody Picture Vocabulary Test—Third Edition, and reading comprehension (Stanford Achievement Test, 10th ed.) tasks multiple times in each year. Students' initial status and rate of growth in each predictor within each grade were estimated using individual growth modeling. These estimates were then used as predictors in dominance regression analyses to examine relative contributions that the predictors made to the outcome: reading comprehension. Among the 1st-grade predictors, individual differences in growth rate in oral reading fluency in 1st grade, followed by vocabulary skills and the autoregressive effect of reading comprehension, made the most contribution to reading comprehension in 3rd grade. Among the 2nd- and 3rd-grade predictors, children's initial status in oral reading fluency had the strongest relationships with their reading comprehension skills in 3rd grade.

Keywords: DIBELS, dominance analysis, growth rate, oral reading fluency, reading comprehension

Since the passing of the No Child Left Behind Act of 2001, the importance of developing students' reading skills, reading fluency in particular, is at the forefront of discussion (e.g., Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008). An underlying assumption in this discussion is that students will have differential growth rates in reading development and that variation will matter for their reading acquisition. In other words, because students vary widely in their initial reading skills, students who enter school with limited language and literacy experiences may need to catch up or develop at a faster rate than their peers with more experience. Therefore, not only overall level of skill but also growth rate in critical reading skills should be predictive of later reading achievement. However, despite the presumed importance of rate of growth, it is not clear whether differences in growth rate in important emergent and conventional literacy skills are predictive of later reading comprehension achievement. Emergent literacy skills refer to "the skills, knowledge, and attitudes that are developmental precursors" (e.g., phonological awareness, letter knowledge, and concept of print; Whitehurst & Lonigan, 1998, p. 848) to conventional forms of reading and writing (e.g., word decoding skills, oral reading fluency [ORF], and spelling).

Examining the contribution of growth rate in emergent and conventional literacy skills to reading comprehension is important for both policy and instructional purposes. From a policy perspective, if variability in growth rate in literacy skills is related to students' later reading comprehension achievement, it is important to incorporate growth rate into accountability decisions. From an instructional perspective, if the growth rate in literacy skills reliably predicts later reading comprehension, teachers may need to frequently monitor growth in order to accommodate to students' changing needs. However, if overall level of student performance at the beginning of the year plays a critical role in determining students' achievement at the end of the year or at a later time point, then teachers can primarily utilize assessment information at the beginning of the year to plan instruction for the year. In the present study, we investigated whether and to what extent level of performance at the beginning of the year (initial status) and growth over time in important literacy skills (e.g., vocabulary, phonological decoding fluency, and ORF) in first, second, and third grades make a contribution to reading comprehension at the end of each grade and third grade.

Background and Context

In the present study, we used data from a statewide database that includes scores from the Dynamic Indicators of Basic Early Literacy Skills assessments (DIBELS; Good & Kaminski, 2002).

Young-Suk Kim and Barbara Foorman, College of Education, Florida State University and Florida Center for Reading Research, Tallahassee, Florida; Yaacov Petscher, Florida Center for Reading Research; Christopher Schatschneider, Department of Psychology, Florida State University and Florida Center for Reading Research.

Correspondence concerning this article should be addressed to Young-Suk Kim, C234D, Florida Center for Reading Research, Department of Psychology, Florida State University, 1107 West Call Street, Tallahassee, FL 32306. E-mail: ykim@fcrr.org

DIBELS was widely used in the primary-grade assessment component of No Child Left Behind Act of 2001, Reading First. DIBELS offers several subtests that measure critical emergent and conventional reading skills such as phonological awareness (e.g., Phoneme Segmentation Fluency subtest), phonological decoding skills (Nonsense Word Fluency subtest), and ORF. These tests are administered to identify students at risk for reading difficulty and in need of intervention, and to monitor their progress. Of the subskills measured in DIBELS, ORF has been widely used as an efficient way of measuring and monitoring students' progress in overall reading. Considerable research has demonstrated that ORF, operationalized as the number of words read correctly per minute (WCPM), is a good indicator of children's overall reading skills development (e.g., L. S. Fuchs, Fuchs, Hosp, & Jenkins, 2001; Jenkins & Jewell, 1993; Stahl & Kuhn, 2002; Yovanoff, Duesbery, Alonzo, & Tindal, 2005; Wiley & Deno, 2005). According to Perfetti's (1985) verbal-efficiency account of reading, efficient word reading releases attentional resources to attend to meaning in text. Thus, students who read dysfluently (i.e., slowly and with great effort) expend their energy on identifying words rather than getting at meaning. Previous research has shown a strong concurrent relationship between ORF and reading comprehension, particularly for students in the primary grades. For example, correlation coefficients ranged from .73 (Cook, 2003) and .76 (Roberts, Good, & Corcoran, 2005) for first graders and from .67 (Good, Simmons, & Kame'enui, 2001) to .70 (Buck & Torgesen, 2003; Roehrig et al., 2008) for third graders. ORF has also been shown to predict students' later reading comprehension achievement, although the strength of the relationship tends to be weaker. For example, ORF at the end of first grade was significantly associated (r = .54) with reading comprehension at the end of second grade (Ridel, 2007). Unless otherwise noted, in this study we use the term ORF as a reading rate determined by the number of WCPM when reading previously unseen grade-level passages connected text.

While informative, these previous studies have been limited to examining point estimates of the relationships between ORF and reading comprehension, but not estimates of the relationship between growth trajectories of ORF and reading comprehension skills, that is, estimates of how students' initial status (or end-of-year status) and growth rate in ORF are related to their reading comprehension achievement. Response to instruction/intervention (RTI) approaches (Compton, Fuchs, Fuchs, & Bryant, 2006; D. Fuchs, Fuchs, McMaster, & Al Otaiba, 2003; Lyon, Shaywitz, & Shaywitz, 2003) suggest identifying students with future reading difficulties based on students' limited responses (i.e., no or slow growth) to evidence-based scientific instruction. Therefore, not only where students begin/end but also students' growth rate would provide important information for classifying students and/or for accountability decisions. In a recent study, Schatschneider, Wagner, and Crawford (2008) examined predictive validity of growth rate compared to end-of-year status, thus examining the utility of growth rate information for accountability purposes. Schatschneider et al. examined whether variability in growth rate in ORF provides unique information on students' reading comprehension achievement above and beyond children's end-of-year performance on ORF. Their results showed that in first-grade, end-of-year status in ORF was positively related to reading comprehension at the end of first and second grades, whereas growth rate in ORF did not make any meaningful contribution to reading comprehension above and beyond end-of-year status. Several states (e.g., Tennessee, Ohio, and North Carolina) have adopted growth rate as a measure of Adequate Yearly Progress (AYP) for No Child Left Behind Act of 2001. However, the actual value of incorporating growth as part of an accountability measure has not been empirically examined. Schatschneider et al.'s study suggests that for making decisions about accountability, variability in growth rate in ORF may not provide any additional information beyond a single time point assessment at the end of the year.

Another important reason to consider examining the relationship between growth trajectories of ORF and reading comprehension is for screening purposes. Similar to the RTI approaches (Compton et al., 2006; D. Fuchs et al., 2003; Lyon et al., 2003), teachers are expected to use assessment results in order to screen and identify struggling readers and provide appropriate and differentiated instruction that is approved based on scientific research. The main goal is for teachers to identify students with potential reading problems early in their reading development in the beginning of the year and also continue to frequently monitor students' progress throughout the year. Although continued monitoring of students' progress is key to any successful instruction (particularly for students who struggle), it is an empirical question whether students' growth rate in several important literacy skills provides additional information about their later reading comprehension achievement above and beyond their initial status in the beginning of the year. On the one hand, it is possible that where students begin in the fall in important predictors most strongly predicts where they end up in the spring. Thus, students' initial status in critical reading skills might provide the most information for identification and instructional purposes for many students. On the other hand, variability in growth rate in critical reading skills may provide additional or more information about students' later reading achievement. For example, the RTI dual discrepancy model advocates use of both students' performance level and growth rate to identify children who do not respond to quality instruction or intervention (i.e., students whose growth rate is slower than that of their classmates) (D. Fuchs et al., 2003; L. S. Fuchs & Fuchs, 1998; McMaster, Fuchs, Fuchs, & Compton, 2005; Speece & Case, 2001). In the present study, we did not examine the efficacy of the dual discrepancy model or overall RTI framework (e.g., frequent progress monitoring for a few nonresponders). Rather, the goal of the present study was to understand the big picture about the extent to which initial status (typically used for screening purposes) as well as growth rate in several emergent and conventional literacy skills might be differentially related to reading comprehension achievement, using statewide benchmark testing data collected four times a year.

In summary, in the present study we examined the following research questions: (a) to what extent initial status and growth rate in emergent and conventional literacy skills measured by DIBELS measures (i.e., letter-naming fluency [LNF], phonological decoding fluency, and ORF) and vocabulary from first grade to third grade are related to reading comprehension skills at the end of each grade (i.e., how initial status and growth rate in each grade is related to reading comprehension at the end of each grade; proximal outcome) and (b) to what extent initial status and growth rate in the same emergent and conventional literacy skills in first grade to third grade are related to reading comprehension in third grade (distal outcome). Third-grade reading comprehension was chosen as the distal outcome in the analysis because end-of-third-grade reading comprehension is important in an accountability system as students are expected to read at grade level by the end of third grade (No Child Left Behind Act of 2001). Furthermore, it was presumed theoretically that ORF is a relatively higher order reading skill such that other lower level, more fundamental skills (i.e., phonological awareness, LNF, and phonological decoding fluency) would be subsumed under ORF (L. S. Fuchs et al., 2001; Katzir, Shaul, Breznitz, & Wolf, 2004). It is an empirical question, however, whether and, if so, how much unique information the lower level skills provide for children's later reading comprehension achievement above and beyond ORF. In practice, students' performance on lower level skills provides teachers with critical information about potential sources of students' struggle in reading. ORF, although useful, does not inform teachers about why students have difficulties with fluent reading. Thus, in addition to ORF, the information on phonological awareness, LNF, phonological decoding fluency, and vocabulary will help determine the nature of instruction and intervention. We addressed these questions using 4-year, large-scale longitudinal data beginning in kindergarten from students attending Reading First schools in Florida. During the Reading First implementation, schools in Florida were required to provide instruction from a core reading program that was aligned with the five key components of reading identified by the National Reading Panel (National Institute of Child Health and Human Development, 2000), and to monitor student progress four times a year.

Method

Participants

A sample of students from Florida's Reading First population were drawn from the Progress Monitoring and Reporting Network that is maintained by the state of Florida. The Progress Monitoring and Reporting Network is an archival database that contains deidentified student information on reading performance for over 1.2 million kindergarten to 12th grade students in Florida. Participants were selected based on their attendance at a Reading First school over a 4-year period (2003–2007), spanning from kindergarten through third grade. From the available 18,556 students who were enrolled in a Reading First school at any point during kindergarten through third grade, we selected students who were consistently enrolled across the 4-year time period (N = 13,154). Approximately 5% of this sample was not tested in reading comprehension skills at the end of third grade, resulting in a total of 12,536 participants retained in the final sample. Students were evenly distributed across gender (50% female), and were diverse in ethnicity with 41% White (n = 5,191), 30% Black (n = 3,781), 22% Latino (n = 2,762), 4% multiracial (n = 552), 2% Asian (n = 201), and less than 1% Native American (n = 49) students. As a proxy for socioeconomic status, the percentage of students eligible for free or reduced-priced lunch was used to describe the degree of economic disadvantage in the sample. Thirty-one percent of the students were not free or reduced-priced lunch eligible, 52% were eligible for reduced-price lunch, and 17% were eligible for free lunch. This sample was representative of 310 out of the population of 318 Reading First schools at the time from which the sample was drawn, and was representative of 33 out of 67 districts, including both rural and urban areas across the state.

Measures

Outcome: Reading comprehension. The Stanford Achievement Test (10th ed.; SAT-10; Harcourt Brace, 2003), a nationally normed test, was used to measure reading comprehension. This test includes several passages followed by multiple-choice items that assess students' initial understanding, interpretation, critical analysis, and awareness and usage of reading strategies after reading literary, informational, and functional text passages. The SAT-10 was administered at the end of each year from first grade to third grade. Reliability for the SAT-10 reading comprehension subtest was reported to be .88 for a nationally representative sample. The SAT-10 manual also reported strong evidence of content, criterion-related, construct, and convergent validity (Harcourt Brace, 2003).

Predictors: DIBELS Subtests and Receptive Vocabulary

Predictors. DIBELS measures included phoneme segmentation fluency, LNF, nonsense word fluency, and ORF. In Florida, during the present study, these DIBELS measures were administered four times a year. The first assessment occurred within the first 20–30 days of school (September), the second between the 65th–75th days of school (December), the third between the 110th–120th days of school (February), and the fourth between the 155th–165th days of school (April). In addition, children's receptive vocabulary was measured at the end of each year (April) from kindergarten to third grade, using the Peabody Picture Vocabulary—Third Edition (PPVT-III; Dunn & Dunn, 1997).

Phonemic Segmentation Fluency (PSF). PSF assesses children's phonological awareness fluency (Kaminski & Good, 1996, 1998). The child is asked to segment the phonemes in an orally presented word containing three or four phonemes. The number of correct segments produced in 1 min is used, allowing partial credit. For example, for the word *sat*, a response of /s/ and /at/ would earn a score of 2 compared to a full credit of 3 at the phoneme level /s/ /a/ /t/. Alternate-form reliability is .90. PSF was administered four times during first grade, and a single form of PSF was used.

LNF. LNF assesses children's ability to name letters (Kaminski & Good, 1996, 1998). An array of upper- and lowercase letters is presented in random order, and students are asked to name as many letters as they can in 1 min. Students are told that if they do not know a letter, the examiner will provide it for them. Alternate-form reliability is .99. LNF was administered only at the beginning of first grade.[1]

Nonsense Word Fluency (NWF). NWF measures children's decoding fluency using nonwords (Kaminski & Good, 1996, 1998). The child is asked to read vowel–consonant and consonant–vowel–consonant, single-syllable nonwords that have short vowels. The child is told to read the make-believe words as quickly and as accurately as possible. If the child does not respond

[1] Letter naming fluency has only 1 data point in first grade. However, it was included in the analysis given its predictive validity for ORF (Speece, Mills, Ritchey, & Hillman, 2003).

within 3 s, the examiner prompts with "next?" The stimuli are presented in 12 rows of five words each. Scoring allows credit for correctly producing individual phonemes or for producing the nonword as a blended unit. For example, if the nonword is *vab*, 3 points are awarded if the child says either /v/ /a/ /b/ or *vab*. Alternate-form reliability is .83. The NWF task was administered four times in first and second grades with a total of 8 data points. A single form of NWF was used.

ORF. DIBELS ORF (5th ed.; Good, Kaminski, Smith, Laimon, & Dill, 2001) assesses ORF for reading grade-level connected text. The child is asked to read three grade-level passages aloud for 1 min, and the number of words accurately read in 1 min is calculated. Words omitted, substituted, and hesitations of more than 3 s are scored as errors. In accordance with the DIBELS guideline, the median score from the three passages was used as the final score. Good, Kaminski, et al. (2001) reported median alternate-form reliability for oral reading of passages to be .94. Although ORF is not typically administered until January of first grade, in Florida ORF was administered from the beginning of first grade to the end of third grade with a total of 12 waves of data.

Receptive vocabulary. Florida Reading First schools also assessed students' receptive vocabulary using the PPVT-III (Dunn & Dunn, 1997) at the end of each year (April). In this task, the child is asked to select from among four pictures the one picture named by the examiner. The median split-half reliability coefficient is .80, and test–retest reliability is .92. The PPVT-III shows a high degree of stability across time, so students' end-of-year scores inform instruction in the following year. Therefore, in our regression analysis, we used children's PPVT-III score from a previous academic year (e.g., end of kindergarten) as a predictor for reading comprehension at the end of the current year (e.g., end of first grade).

Data Analytic Strategies

Multilevel growth modeling. A combination of multilevel modeling and dominance analysis were the primary analytic procedures in the present study. Multilevel growth modeling was used in order to estimate students' average level of performance at the beginning of each grade (i.e., initial status), as well as the average rate of change for each of the predictors (i.e., PSF, NWF, and ORF). Since participants were assessed on DIBELS measures multiple times during the school year, growth over time can be viewed as being nested within each individual (thus multilevel). The estimates obtained from the growth models (i.e., individual variation around the average initial status and growth rate) were then used in a dominance analysis as predictors of reading comprehension outcome at the end of first, second, and third grades. This latter analysis revealed which predictor(s) made significant contributions to explaining variation in the reading comprehension outcome.

The DIBELS subtests vary in the overall frequency and the time points when they are administered from kindergarten through third grade, an important consideration when estimating growth models. For example, the NWF task was administered a total of 10 times from kindergarten to second grade. As the number of time points an assessment is given increases, so does the complexity of describing growth. With three time points, a linear growth model may be used to describe the average trajectory across students, but when four time points or more exist for an individual, nonlinearity may be tested by adding additional growth parameters to the model (e.g., a quadratic term).

A further consideration relevant to the development of growth models in this study was accounting for the growth not only within grade, but across grades as well. The DIBELS tasks were given over multiple grades. By fitting a traditional growth model across multiple years, we found that slope estimates not accounting for the deceleration that typically occurs in learning over the summer may produce biased estimates of within-year growth. Such drop-offs can be accounted for by using piecewise growth curve models (Raudenbush & Bryk, 2001), which allows for the estimation of multiple growth rates within the same model. The piecewise growth curve model is an extension of the more traditional growth curve model, with the exception that it allows the researcher to simultaneously test separate growth profiles within the same regression model (Neter, Kutner, Wasserman, & Nachtsheim, 1996). From this framework, it is expected that the linear function in initial status and growth within each grade would be more homogeneous, but heterogeneous across grades.

The increased complexity of the piecewise growth curve model can be seen from the formulaic representation in Appendix A. Growth is modeled for student i at time t, where $t = 0$ for the month of September. The growth model for one piece in time includes the intercept (π_{0i}), a linear growth parameter (π_{1i}), a quadratic parameter that indicates the rate of celeration (π_{2i}; i.e., acceleration or deceleration), and the error term e_{ti}. Subsequently, π_{0i}, π_{1i}, and π_{2i} in the current model represent the intercept, linear, and quadratic pieces, respectively, for first graders and act as a referent group. Designation of the piecewise components of the model are indicated by $\pi_{3i} - \pi_{8i}$ in the Level 1 model, which are the second- and third-grade intercepts and slopes. These covariates represent dummy coded parameters that are interpreted as fitted deflections from the first-grade components of the model.

In order to select the most valid model to describe growth across tasks, we estimated a series of models that contained a combination of fixed and random parameters. Fixed parameters were the mean predicted initial status scores (i.e., September scores) and growth rate across the students, while random parameters were the estimation of variability around the mean (i.e., variance components). The five models that were examined included (a) linear growth model with a random intercept and a fixed slope; (b) linear growth model with a random intercept and a random slope; (c) nonlinear growth model with random intercept, fixed linear slope, fixed quadratic slope; (d) nonlinear growth model with random intercept, random linear slope, fixed quadratic slope; and (e) nonlinear growth model with random intercept, random linear slope, random quadratic slope. Each model was tested across the different measures. For measures that were administered across multiple years, the piecewise model was applied, and the five models were tested for each piece. Using the outlined procedures, we used Model d to describe second- and third-grade growth in ORF, and we used Model e to describe growth in the first- and second-grade pieces of the NWF task, the first-grade piece of the PSF task, and the first-grade piece of the ORF task. Models a, b, and c were not used to fit any of the data because they were not the best fitting model for our data.

Each of the different models was estimated using Hierarchical Linear Modeling 6.06 (Raudenbush, Bryk, & Congdon, 2004), and the individual ordinary least squares growth rates for each student were retained. The individual student slope from the linear model was the amount of growth per month students made for the given task for the grade the model estimated. Similarly, growth from the quadratic model was measured by both a linear component and a rate of celeration (i.e., acceleration or deceleration). The linear component represented the amount of growth, while the quadratic component represented the amount of change in the slope during the course of the year. As previously described, individual students' initial status and growth rates from these models were estimated and subsequently used in a dominance analysis to predict reading comprehension scores on the SAT-10 at the end of the year in Grades 1–3.

Dominance analysis. Many researchers who have examined the concurrent or predictive utility of oral reading measures have used stepwise or hierarchical regression techniques in order to evaluate the unique contribution of a predictor to a model that includes other correlated variables. Both stepwise and hierarchical regression are special cases of the multiple regression, whereby a series of predictors may be entered into an equation with a direct relationship being estimated, controlling for the effect of other predictors in the model. A particular difficulty with using such methods in the social sciences to estimate unique variance is that predictors are typically correlated with each other. A hierarchical multiple regression assumes that the order of entry for each predictor is theoretical; specifically, with a group of multiple independent variables, the order with which each is added to the model affects the estimate of the unique variance component. More limiting than the hierarchical regression is the use of a stepwise regression, which is widely denigrated as an inefficient and often inappropriate method to estimate unique variances since it relies on statistical software-based decisions, rather than using empirical evidence to guide the analysis (Thompson, 1995).

The problem of assessing importance becomes more difficult as predictors become more correlated. Budescu (1993) developed a method of multiple regression, called dominance analysis, that addresses the nature of correlated predictors. Dominance analysis involves the pairwise comparisons of all predictors as they relate to a criterion (Budescu, 1993). Predictor A (e.g., growth rate of ORF) is considered more important, or dominant, over Predictor B (e.g., initial status in ORF) in predicting the criterion (e.g., year-end reading comprehension) if the unique variance contribution to the criterion of Predictor A is greater than Predictor B when both are simultaneously in the model without other predictors, and when both are in the presence of other predictors in the model.

We examined all possible model combinations of predictors to determine an order of importance. Using 2^{p-1} model, where p is the number of predictors based on the number of variables one wishes to include in the analysis, one is able to calculate the total number of subset models for the dominance analysis. In the context of the current study, a maximum of five predictors was used (see Tables 2, 3, and 4), which resulted in 31 total subset models. The subset models comprised a series of multiple regressions where pairwise comparisons of predictors are entered. Thus, in the example of five predictors, there are five subset models with only one predictor, 10 subset models with a combination of two predictors, 10 subset models with a combination of three predictors, five subset models with four predictors, and one subset model with all five predictors, producing 31 total subset models.

In addition to estimating the pairwise comparisons among unique variance contributions, the difference in the R^2 between predictors (ΔR^2) was tested for statistical significant. Since the R^2 was estimated within the same sample, a formula provided by Alf and Graf (1999) was used to calculate the asymptotic standard error for each comparison. Using ΔR^2 and the standard error, we constructed 95% confidence intervals to evaluate whether the differences were statistically significant. Intervals not containing zero were statistically significant at the 0.05 level. A ΔR^2 value of 0.04 was considered to be a small but practically important difference, 0.25 was identified as a moderate difference, and 0.64 was identified as a large difference (Ferguson, 2009).

Using these analytical strategies, in this study we aimed to determine the extent to which initial (or end-of-year) status or growth rate in emergent and conventional literacy skills in first grade to third grade (i.e., phonological awareness, phonological decoding skills, and ORF) predict proximal (reading comprehension at the end of first, second, and third grades) and distal outcomes (i.e., reading comprehension in third grade). Additionally, it was of interest to examine how ORF growth in first and second grades compares to either vocabulary or reading comprehension in first and second grades as a predictor of third-grade reading comprehension, given the well-established role of vocabulary in reading comprehension (e.g., National Institute of Child Health and Human Development, 2000), and the theoretical importance (L. S. Fuchs et al., 2001; Perfetti, 1985) and empirical evidence (e.g., Ridel, 2007; Roehrig et al., 2008) of the role of reading fluency in reading comprehension.

Results

Preliminary Analysis

Means and standard deviations for each measure by wave and grade are presented in Table 1. Keeping in mind that our sample was drawn from Reading First schools, it is noteworthy that in each grade and at each time point, mean NWF, Phonemic Segmentation, and ORF were well above the state fall and spring benchmarks (Good & Kaminski, 2002). For example, for NWF in first grade, "grade-level" performance is above 24 WCPM in the fall and 50 WCPM or greater in the spring. A similar picture was observed for ORF (e.g., for first grade, grade-level performance is above 7 WCPM in the fall and 40 WCPM or greater in the spring). Specifically, when the students were in first grade, their mean ORF scores at both the beginning of the year (M = 22 WCPM) and end of the year (M = 56 WCPM) were well above the fall and spring benchmarks (i.e., ≥ 7 WCPM and ≥ 40 WCPM, respectively). Similarly, during second grade, students' mean fall ORF scores (M = 59 WCPM) were at the grade-level benchmark of ≥13 WCPM, as were their spring ORF scores (M = 97 WCPM, benchmark ≥ 90 WCPM). Third-grade fluency scores followed the same pattern with students' mean fall ORF scores (M = 87 WCPM) meeting the benchmark estimate. Student performance on the vocabulary test indicated that while they performed slightly below a normative mean of 100 in first grade (M = 95.45), by third grade their performance (M = 100.28) had caught up relative to a normative national peer group.

Table 1
Means and Standard Deviations for DIBELS and Outcomes Measures

Grade and assessment	LNF	PSF	NWF	ORF	PPVT-III	SAT-10
First						
Fall	50.6 (15.1)	41.1 (16.3)	38.7 (21.6)	22.5 (19.9)		
Winter 1		45.3 (13.6)	54.4 (24.9)	35.1 (25.6)		
Winter 2		51.5 (12.7)	61.1 (28.3)	45.5 (30.1)		
Spring		48.6 (12.4)	71.5 (32.1)	55.5 (30.4)	95.5 (14.2)	561.3 (45.9)
Second						
Fall			68.5 (31.8)	59.2 (31.1)		
Winter 1			84.2 (37.5)	72.8 (30.1)		
Winter 2			90.7 (38.7)	85.8 (33.6)		
Spring			99.1 (42.3)	96.9 (33.6)	98.5 (14.0)	604.1 (37.6)
Third						
Fall				86.1 (30.6)		
Winter				101.1 (31.3)		
Spring				121.3 (34.5)	100.3 (13.3)	632.6 (41.4)

Note. Standard deviations are shown in parentheses. DIBELS = Dynamic Indicators of Basic Early Literacy Skills; LNF = Letter-Naming Fluency; PSF = Phonemic Segmentation Fluency; NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

Estimates of Growth

As noted above, we obtained fitted initial status and growth rate estimates of predictors for each student by using multilevel growth models. Figure 1 displays the mean trajectories of NWF and ORF for each grade in the piecewise model, and Appendix B reports the estimated monthly growth rates. Results from the growth model on ORF indicated that when students were in first grade they grew, on average, 3 WCPM per month. The associated estimated rate of celeration (i.e., the amount of acceleration or deceleration in slope) in first grade was 0.25, which suggested that over time students grew faster. Based on the first-grade rate of celeration, students' growth rate in ORF doubled from 3 WCPM in September to gains of 6 WCPM in March. During second grade, the estimated linear growth rate was 4.4 WCPM with a celeration rate of 0.15. The relationship between these two was such that the increase of growth went from 4.4 WCPM in September to 6.2 WCPM in March. Lastly, the average growth in third grade was 5.6 WCPM per month. Because the quadratic parameter was not statistically significant in the model, this term was dropped and only the linear rate was used to calculate growth. Results for NWF indicated that when students were in first grade, they grew, on average, 5 WCPM in first grade, but there was slight deceleration (i.e., −0.04) such that the average monthly growth rate was 4.29 in March. A similar pattern was observed in second grade with a larger deceleration rate.

Interrelationships among the predictors (including estimated initial status and growth rates obtained from the above individual growth modeling) and criterion variables for first, second, and third grade[2] are reported in Table 2. In first grade, growth in ORF (ORFG) was more strongly correlated with third-grade scores on the SAT-10 (r = .64) than either ORF from the fall (r = .49) or spring (r = .59). Moreover, first-grade ORFG was nearly as correlated with end-of-first-grade SAT-10 performance (r = .71) as was the spring assessment of ORF (r = .73). The large positive correlation between the fall ORF administration and monthly ORFG was an indication that students who started the first grade with a higher fluency rate tended to grow faster than those with a lower fluency rate.

Conversely, in both second and third grades, the fall administration of ORF was more strongly related to SAT-10 in third grade (r = .62 and r = .66, respectively) than were estimates of ORFG within each grade (r = −.44 and r = −.38, respectively). The moderate negative correlations indicated that students who had high levels of ORF in the fall had slower growth during their respective grades.

Dominance Analysis

We conducted a series of five dominance analyses in order to examine the relative dominance of the predictors within each grade: (a) using first-grade student performance to predict end-of-year first grade SAT-10, (b) using first-grade student performance to predict end-of-year third-grade SAT-10, (c) using second-grade student performance to predict end-of-year second-grade SAT-10, (d) using second-grade student performance to predict end-of-year third-grade SAT-10, and (e) using third-grade student performance to predict end-of-year third-grade SAT-10. Since third grade's end-of-year test serves as the distal outcome for the first- and second-grade analyses, only one dominance analysis was done in third grade. As stated above, a maximum of five predictors were included in the dominance analyses. The predictors were selected based on their bivariate relationships with reading comprehension in each grade. For example, growth in PSF and NWF in first grade was not related to reading comprehension (r < .10). Thus, the following five predictors were included when first-grade SAT-10 was predicted by first-grade predictors (see Table 3): fall scores of LNF, NWF, and ORF; ORFG; and PPVT-III score at the end of kindergarten. Results are first reported for the prediction of within-year reading comprehension, and then for the prediction of distal

[2] Correlations among all the observed variables by wave and grade can be obtained from the first author upon request.

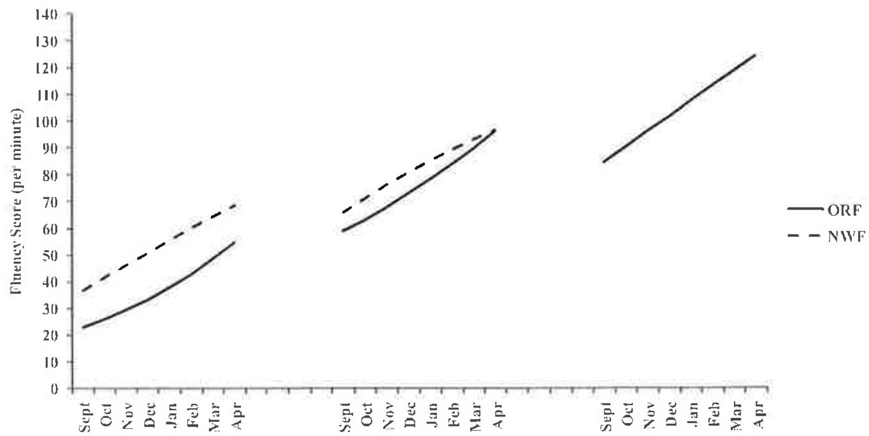

Figure 1. Predicted growth trajectories of oral reading fluency and nonsense word fluency (words correct per minute) in first, second, and third grades.

outcome (i.e., reading comprehension at the end of third grade). The tables for the dominance analyses are organized such that the first column labeled *Subset model* presents the independent variables that were entered into the multiple regression, the second column contains the total R^2 value for that subset model, and remaining columns report the unique variance contribution added to that subset model. For example, in Table 3, the subset model of LNF demonstrates that 24% of the variance in first-grade SAT-10 scores was accounted for by LNF. When controlling for LNF, first-grade ORFG contributed 27% unique variance, fall ORF scores uniquely added 17% variance, fall NWF added 10% variance, and students' kindergarten vocabulary added 7% unique variance above LNF. As the subset models increase in the number of predictors, the R^2 values represent the variance jointly accounted for in the outcome. In the subset model of LNF–NWF in Table 3, 34% of the variance was accounted for by both predictors (i.e., LNF and NWF), with 19% unique variance added by ORFG, 8% added by fall ORF, and 6% added by PPVT-III in kindergarten.

While these tables represent the estimation of unique and total variance components from iterative multiple regression models, the more important component was the statistical test of differences in the unique R^2 values. Tables 5 and 9 report the tests of ORFG by grade compared to other within-grade predictors. The first column in these tables represent the pairwise comparison that was made, with the second column reporting the difference in the R^2 values, followed by the standard error, and the 95% confidence interval for the R^2 difference. The comparisons within this table first look at the difference in unique variance added when the predictors are included in the subset model together with a unique variance added when each individual predictor from that subset model is added to the other. As an example, for the comparison of unique variance contributions between ORFG and LNF in first grade (see Table 5), we considered the additional contribution of variance of each predictor to the other when only examining models with one predictor. In this case, the unique contribution of ORFG to a model with only LNF is .27, which is obtained from the LNF–ORFG subset model R^2 of .51 minus the LNF subset model R^2 of .24 (see Table 3). The unique contribution of LNF to a model with only ORFG is .01 and is obtained by taking the LNF–ORFG subset model R^2 of .51 minus the ORFG subset model R^2 of .50 (see Table 3). The difference of these two unique variance components of .50 and .24 is .26, and is reflected in Table 5. With the large sample size, the observed standard error was small ($SE = 0.004$), and the 95% CI was [0.252, 0.268]. Because this interval did not contain 0, we can conclude that the difference in unique variance contributed by the two predictors is statistically significant.

When determining the relative importance between LNF and ORFG (i.e., ORFG) in the presence of the other predictors, the method for obtaining the ΔR^2 occurs by examining the amount of unique variance each predictor adds to a model that contains all other variables, and taking the difference between those two estimates. Thus, in Table 3, the amount of variance that ORFG adds to the model with all other predictors (i.e., LNF–NWF–ORF–PPVT-III) was 0.13, and the amount of variance that LNF adds to its respective model (i.e., NWF–ORF–ORFG–PPVT-III) is 0. The difference of these two unique estimates (i.e., $0.13 - 0.00$) is reflected in Table 5. The standard error for this contrast was also 0.004, with an associated 95% CI of [0.123, 0.137]. Therefore, in this model, when fall ORF, fall NWF, and kindergarten PPVT-III are controlled, ORFG contributes more unique variance in predicting end-of-first-grade reading comprehension than does LNF.

Reading Comprehension at the End of First Grade

A summary of the results for the dominance analysis and pairwise difference tests for first-grade reading comprehension outcome are reported in Tables 3 and 5, respectively. Examination of the unique variance components in Table 3 indicates that ORFG in first grade was a dominant predictor of first-grade SAT-10 scores. For each occasion that ORFG was uniquely added to the subset model, its unique contribution was greater than all other predictors. On average, ORFG uniquely contributed 24.5% variance across all subset models, followed by fall ORF (12.9%), fall NWF (8.4%), fall LNF (6.3%), and kindergarten vocabulary (6.2%). When all of the predictors except for ORFG were in the model, 45% of the variance was accounted for, yet growth still uniquely contributed 13% to explaining the variance. This resulted in a total of 58% of reading comprehension performance explained by all five predic-

Table 2
Correlations Among DIBELS, PPVT-III, and SAT-10 in Grades 1–3

	First grade										Second grade								Third grade			
Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
										First grade												
1. LNF	—																					
2. NWF	.62	—																				
3. PSF	.42	.44	—																			
4. ORFF	.54	.72	.3	—																		
5. ORFS	.58	.69	.33	.80	—																	
6. ORFG	.56	.72	.31	.59	.90	—																
7. NWFG	−.46	−.52	−.37	−.51	−.51	−.46	—															
8. PSFG	−.36	−.34	−.74	−.32	−.30	−.25	−.37	—														
9. PPVT-III	.29	.27	.35	.32	.31	.30	−.21	−.28	—													
10. SAT-10	.49	.60	.37	.61	.73	.71	−.33	−.34	.40	—												
										Second grade												
11. NWF	.66	.80	.28	.78	.74	.63	−.42	−.17	.30	.59	—											
12. ORFF	.59	.70	.32	.76	.91	.94	−.48	−.29	.32	.73	.71	—										
13. ORFS	.56	.67	.31	.63	.83	.93	−.46	−.26	.30	.66	.64	.87	—									
14. ORFG	−.49	−.68	−.27	−.72	−.89	−.76	−.36	.28	−.26	−.67	−.66	−.87	−.14	—								
15. NWFG	−.20	−.14	−.08	−.29	−.12	−.01	.03	.10	−.11	−.11	−.39	−.12	.00	.19	—							
16. PPVT-III	.29	.29	.35	.34	.34	.32	−.22	−.28	.77	.43	.31	.34	.33	−.28	−.10	—						
17. SAT-10	.41	.49	.36	.48	.57	.61	−.26	−.31	.47	.64	.48	.60	.63	−.43	−.07	.50	—					
										Third grade												
18. ORFF	.55	.66	.29	.64	.83	.92	−.46	−.25	.31	.66	.63	.87	.91	.04	−.02	.33	.63	—				
19. ORFS	.06	.07	.05	.06	.09	.10	−.06	−.03	.09	.10	.07	.09	.10	−.06	.01	.09	.10	.10	—			
20. ORFG	−.42	−.58	−.23	−.59	−.76	−.75	.37	.23	−.22	−.60	−.53	−.77	−.65	.56	.08	−.24	−.42	−.25	−.06	—		
21. PPVT	.30	.35	.32	.39	.41	.40	−.25	−.27	.68	.47	.36	.42	.41	−.33	−.08	.73	.54	.43	.11	−.30	—	
22. SAT-10	.40	.44	.19	.49	.59	.64	−.12	−.18	.47	.63	.47	.62	.62	−.44	−.06	.51	.73	.66	.10	−.38	.58	—

Note. All coefficients are statistically significant at the .05 level. The letter *F* after a variable name represents fall performance; the letter *S* represents spring performance. NWFG and ORFG represent growth rates in NWF and ORF, respectively. DIBELS = Dynamic Indicators of Basic Early Literacy Skills; LNF = Letter-Naming Fluency; PSF = Phonemic Segmentation Fluency; NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

Table 3
Dominance Analysis Results From First-Grade Variables Predicting SAT-10 at the End of First Grade

		Unique contribution of predictor to first-grade SAT-10				
Subset model	R^2	LNF	NWF	ORF	ORFG	PPVT-III
Models with one predictor						
LNF	.24		0.10	0.17	0.27	0.07
NWF	.30	0.03		0.09	0.23	0.07
ORF	.37	0.04	0.02		0.18	0.05
ORFG	.50	0.01	0.03	0.06		0.04
PPVT-III	.16	0.16	0.21	0.26	0.38	
Models with two predictors						
LNF–NWF	.34			0.08	0.19	0.06
LNF–ORF	.41		0.00		0.15	0.04
LNF–ORFG	.51		0.02	0.05		0.03
LNF–PPVT-III	.31		0.08	0.13	0.23	
NWF–ORF	.39	0.02			0.16	0.04
NWF–ORFG	.53	0.00		0.03		0.03
NWF–PPVT-III	.37	0.02		0.07	0.19	
ORF–ORFG	.56	0.00	0.00			0.03
ORF–PPVT-III	.42	0.03	0.02		0.16	
ORFG–PPVT-III	.54	0.01	0.02	0.04		
Models with three predictors						
LNF–NWF–ORF	.41				0.14	0.04
LNF–NWF–ORFG	.53			0.03		0.03
LNF–NWF–PPVT-III	.39			0.06	0.17	
LNF–ORF–ORFG	.56		0.00			0.02
LNF–ORF–PPVT-III	.45		0.00		0.14	
LNF–ORFG–PPVT-III	.55		0.02	0.04		
NWF–ORF–ORFG	.56	0.00				0.03
NWF–ORF–PPVT-III	.44	0.01			0.15	
NWF–ORFG–PPVT-III	.56	0.00		0.02		
ORF–ORFG–PPVT-III	.58	0.00	0.00			
Models with four predictors						
LNF–NWF–ORF–ORFG	.56					0.02
LNF–NWF–ORF–PPVT-III	.45				0.13	
LNF–NWF–ORFG–PPVT-III	.56			0.02		
LNF–ORF–ORFG–PPVT-III	.58		0.00			
NWF–ORF–ORFG–PPVT-III	.58	0.00				
Model with five predictors						
LNF–NWF–ORF–ORFG–PPVT-III	.58					

Note. ORFG represents ORF growth. Others represent initial status in each variable (i.e., fall performance) except PPVT-III. LNF = Letter-Naming Fluency; NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

tors. Fall ORF tended to show conditional dominance, meaning that only under certain circumstances does the predictor dominate others in unique variance contribution. In the current context, fall ORF only dominates other predictors when ORFG is already included in the subset model. The remaining three predictors in first grade (LNF, NWF, and PPVT-III) were contributing unique variance at similar levels, and none of the three dominated the others. As noted in Table 5, the differences in the amount of unique variance contributed by ORFG and the other predictors were statistically significant. Moreover, not only were the main contributions of ORFG significant, but in the presence of other variables ORFG contributed a statistically significant greater amount of unique variance.

Reading Comprehension at the End of Second Grade

Results for the prediction of second-grade reading comprehension outcome are provided in Tables 4 and 5. Unlike the first-grade findings, ORFG in second grade was not a dominant predictor of reading comprehension at the end of second grade. On average, it only accounted for 7.1% of the variance across the subset models. When comparing the unique contributions between ORFG and the other predictors, the results of the difference in pairwise comparison indicated that ORFG contributed significantly less variance than all other predictors, except for growth in NWF. Conversely, fall ORF was the strongest predictor of reading comprehension scores, uniquely contributing, on average, 18.6% of the variance. This was followed by end-of-first-grade vocabulary (14.1%), fall NWF (8.2%), ORFG (7.1%), and NWF growth (.36%). The amount of unique variance contributed by fall ORF was greater than unique components of the other variables, with the exception of two subset models. In the model where fall NWF was solely entered, and entered with NWF growth, vocabulary either contributed more unique variance (13% vs. fall ORF, 11% in the NWF fall–NWF growth model) or the same amount (13% in the NWF

Table 4
Dominance Analysis Results From Second-Grade Variables Predicting SAT-10 at the End of Second Grade

Subset model	R^2	Unique contribution of predictor to second grade SAT-10				
		ORF	ORFG	NWF	NWFG	PPVT-III
Models with one predictor						
ORF	.35		0.03	0.01	0.00	0.10
ORFG	.19	0.20		0.07	0.00	0.15
NWF	.23	0.13	0.02		0.02	0.13
NWFG	.01	0.35	0.18	0.24		0.24
PPVT-III	.25	0.21	0.09	0.12	0.00	
Models with two predictors						
ORF–ORFG	.38			0.01	0.00	0.09
ORF–NWF	.36		0.03		0.00	0.09
ORF–NWFG	.35		0.03	0.01		0.10
ORF–PPVT-III	.45		0.03	0.00	0.00	
ORFG–NWF	.26	0.14			0.01	0.12
ORFG–NWFG	.19	0.20		0.08		0.15
ORFG–PPVT-III	.34	0.14		0.04	0.00	
NWF–NWFG	.25	0.11	0.02			0.13
NWF–PPVT	.37	0.09	0.01		0.01	
NWFG–PPVT-III	.25	0.21	0.09	0.13		
Models with three predictors						
ORF–ORFG–NWF	.40				0.00	0.09
ORF–ORFG–NWFG	.39			0.01		0.09
ORF–ORFG–PPVT-III	.48			0.01	0.00	
ORF–NWF–NWFG	.36		0.03			0.09
ORF–NWF–PPVT-III	.45		0.03		0.00	
ORF–NWFG–PPVT-III	.45		0.03	0.00		
ORFG–NWF–NWFG	.27	0.13				0.12
ORFG–NWF–PPVT-III	.38	0.11			0.01	
ORFG–NWFG–PPVT-III	.34	0.14		0.05		
NWF–NWFG–PPVT-III	.38	0.08	0.01			
Models with four predictors						
ORF–ORFG–NWF–NWFG	.40					0.09
ORF–ORFG–NWF–PPVT-III	.48				0.00	
ORF–ORFG–NWFG–PPVT-III	.48			0.01		
ORF–NWF–NWFG–PPVT-III	.46		0.03			
ORFG–NWF–NWFG–PPVT-III	.39	0.09				
Model with five predictors						
ORF–ORFG–NWF–NWFG–PPVT-III	.48					

Note. ORFG and NWFG represents growth in ORF and NWF, respectively. Others represent initial status in each variable (i.e., fall performance) except PPVT-III. LNF = Letter-Naming Fluency; NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

model). Furthermore, in the models where fall ORF was not present, vocabulary was the most dominant predictor.

Reading Comprehension at the End of Third Grade

Prediction of the distal third-grade reading comprehension scores from first, second, and third grades are summarized in Tables 6, 7, 8, and 9. Since the purpose of this research question for first and second grade was to determine the best predictors to use as part of screening for distal difficulties in reading (i.e., reading comprehension in third grade), a different set of predictors were used. Instead of solely using fall scores and ORFG from DIBELS, we opted to add the spring assessment of ORF, as well as using the SAT-10 reading comprehension scale score for that year and the vocabulary score for that year (as opposed to the prior year used in the previous analyses). In other words, when we used first-grade predictors for third-grade reading comprehension performance, we used the following first-grade variables as predictors: fall ORF scores, spring ORF scores, ORFG across first grade, and SAT-10 and PPVT-III at the end of first grade. The same principle applied to predicting third-grade reading comprehension using second-grade predictors (e.g., fall and spring ORF in second grade, ORFG in second grade, and SAT-10 and PPVT-III at the end of second grade). Finally, when third-grade predictors were used to predict reading comprehension at the end of third grade, multiple measures were not available since only ORF is given from the DIBELS test. Dissimilar from the previous grades, the spring ORF assessment was not included as a covariate, because the administration of that test occurs after the SAT-10 is given in Florida. To that end, fall ORF, ORFG, and third-grade vocabulary were utilized in the subset models.

Results from the first-grade prediction of third-grade comprehension revealed that none of the five predictors held complete dominance over the others in unique variance contribution (see Table 6). Generally, ORFG contributed the most unique variance

Table 5
95% Confidence Intervals for Pairwise Differences Predicting First-Grade and Second-Grade SAT-10

Variables compared	ΔR^2	*SE*	95% CI Lower limit	95% CI Upper limit
First grade				
ORFG–LNF	.26	0.004	0.252	0.268
All	.13	0.004	0.123	0.137
ORFG–NWF	.20	0.003	0.193	0.207
All	.13	0.004	0.123	0.137
ORFG–ORF	.12	0.002	0.116	0.124
All	.11	0.003	0.104	0.116
ORFG–PPVT-III	.34	0.005	0.329	0.350
All	.11	0.003	0.104	0.116
Second grade				
ORFG–ORF	−.17	0.003	−0.176	−0.164
All	−.07	0.002	−0.075	−0.065
ORFG–NWF	−.05	0.001	−0.052	−0.048
All	−.02	0.001	−0.023	−0.018
ORFG–NWFG	.18	0.003	0.174	0.186
All	.03	0.001	−0.023	−0.018
ORFG–PPVT-III	−.06	0.001	−0.062	−0.058
All	−.04	0.002	−0.064	−0.056

Note. ORFG and NWFG represents growth in ORF and NWF, respectively. Others represent initial status in each variable (i.e., fall performance) except PPVT-III. LNF = Letter-Naming Fluency; NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

on average (15.1%), followed by first-grade SAT-10 (13.2%), first-grade PPVT-III (11.9%), spring ORF (9.6%), and fall ORF (5.9%). Although the difference in the unique contribution between ORFG and SAT-10 scores was very small (ΔR^2 = .01; see Table 9), the confidence interval for that difference did not include 0, indicating that this was a statistically significant difference. Similarly, the differences between ORFG and the other predictors showed that growth was significantly predicting more unique variance than the other variables.

Replicating findings from the previous second-grade analyses, when predicting third-grade reading comprehension, ORFG in second grade was a weak contributor of unique variance. With the differing set of independent variables in the subset models for the third-grade prediction, growth was the weakest of all predictors and contributed significantly less unique variance than all other second-grade variables in the model (see Table 7). Results indicated that the end-of-year reading comprehension performance in second grade was completely dominant compared to the other predictors. On average, SAT-10 in second grade explained 22.6% of the variance in third-grade reading comprehension, followed by spring ORF (12.8%), second-grade vocabulary (11.7%), fall ORF (11.4%), and ORFG (5.0%). When SAT-10 in second grade was present in the subset model, the remaining variables tended to contribute a similar level of variance, indicating that only SAT-10 in second grade appeared to be a dominant predictor.

Results using third-grade predictors indicated that fall ORF was a dominant predictor of third-grade SAT-10, with an average unique variance contribution of 28.4%. This was followed by PPVT-III (20.5%) and ORFG (5.8%). Comparisons of unique variance between growth, and fall ORF and PPVT-III indicated that significantly less was uniquely contributed by growth than either the other two variables (see Table 9).

Discussion

The goal of the present study was to examine relative contributions of students' initial status and growth rate in several emergent and conventional literacy skills to later reading comprehension achievement using state-wide data from Reading First schools in Florida. The finding that ORF, either initial status or growth rate, was dominant over other emergent and conventional literacy skills (i.e., phonological awareness, LNF, and phonological decoding fluency) provides some support for a hypothesis that ORF is a higher order skill that requires integration of lower level skills (L. S. Fuchs et al., 2001; Katzir et al., 2004). It has been hypothesized that ORF (i.e., oral reading rate of connected text) is a strong predictor of reading comprehension for elementary students because ORF builds upon, and requires integration of, word reading (and lower level skills that contribute to word reading such as phonological awareness and letter knowledge) and postlexical processing in connected text (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003; Wolf & Katzir-Cohen, 2001). Thus, once ORF is accounted for, children's accuracy and fluency in phonological awareness, letter names, and decoding words are not expected to make additional contributions to reading comprehension.

The overall results of the present study suggest that individual differences in ORF growth rate in first grade provide the most information about proximal (i.e., first-grade) and distal (i.e., third-grade) reading comprehension achievement. That is, when we used first-grade predictors, ORF growth rate explained the largest amount of variation in students' reading comprehension at the end of first and third grades (for third-grade reading comprehension outcome, first-grade reading comprehension also made a large contribution). In contrast, when we used second-grade predictors, students' initial status in ORF (i.e., fall performance), not growth rate, provided the most information about the second-grade reading comprehension, which, in turn, was most strongly related to reading comprehension in third grade. Among third-grade predictors, students' ORF in the beginning of the year was more strongly related to year-end reading comprehension than was growth of ORF during third grade. Because our primary interest was to examine whether growth rates of ORF in first, second, and third grades were related to reading comprehension achievement in third grade, we did not examine how ORF growth in first grade was related to second-grade reading comprehension achievement.

The finding that variation in ORF growth rate in first grade was the best predictor of students' later reading comprehension skills in first and third grades underscores the fact that ORF growth rate may provide important information about later reading comprehension achievement. In other words, in addition to students' level in the beginning of the year, it may be important to monitor how fast students improve their ORF, particularly in first grade, in order to identify students who may be at risk of future problems in reading comprehension. The importance of growth rate in first grade may be attributed to the fact that differentiation in ORF growth rate occurred in first grade and this growth rate in first grade highly influenced where students began in second and third grades. In other words, the

Table 6
Dominance Analysis Results From First-Grade Prediction of Third-Grade SAT-10

Subset model	R^2	Unique contribution of predictor to third-grade SAT-10				
		ORFF	ORFS	ORFG	SAT-10	PPVT-III
Models with one predictor						
ORFF	.24		0.11	0.19	0.18	0.14
ORFS	.35	0.00		0.06	0.09	0.11
ORFG	.41	0.02	0.00		0.07	0.11
SAT-10	.39	0.02	0.03	0.08		0.07
PPVT-III	.26	0.11	0.20	0.26	0.21	
Models with two predictors						
ORFF–ORFS	.35			0.09	0.09	0.11
ORFF–ORFG	.43		0.01		0.05	0.10
ORFF–SAT-10	.42		0.02	0.06		0.07
ORFF–PPVT-III	.37		0.09	0.15	0.11	
ORFS–ORFG	.41	0.03			0.07	0.11
ORFS–SAT-10	.43	0.00		0.04		0.07
ORFS–PPVT-III	.46	0.00		0.06	0.04	
ORFG–SAT-10	.48	0.00	0.00			0.07
ORFG–PPVT-III	.52	0.01	0.00		0.03	
SAT-10-PPVT-III	.47	0.01	0.03	0.07		
Models with three predictors						
ORFF–ORFS–ORFG	.44				0.06	0.09
ORFF–ORFS–SAT-10	.43			0.06		0.07
ORFF–ORFS–PPVT-III	.46			0.07	0.04	
ORFF–ORFG–SAT-10	.48		0.02			0.07
ORFF–ORFG–PPVT-III	.52		0.01		0.02	
ORFF–SAT-10–PPVT-III	.48		0.02	0.06		
ORFS–ORFG–SAT-10	.48	0.02				0.07
ORFS–ORFG–PPVT-III	.52	0.01			0.03	
ORFS–SAT-10–PPVT-III	.50	0.00		0.05		
ORFG–SAT-10–PPVT-III	.54	0.00	0.00			
Models with four predictors						
ORFF–ORFS–ORFG–SAT-10	.49					0.06
ORFF–ORFS–ORFG–PPVT-III	.53				0.03	
ORFF–ORFS–SAT-10–PPVT-III	.50			0.05		
ORFF–ORFG–SAT-10–PPVT-III	.55		0.01			
ORFS–ORFG–SAT-10–PPVT-III	.55	0.01				
Model with five predictors						
ORFF–ORFS–ORFG–SAT-10–PPVT-III	.56					

Note. The letters *F*, *S*, and *G* after a variable name represent fall, spring, and growth, respectively. LNF = Letter-Naming Fluency; NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

large correlations between ORF growth rate in first grade and initial status in ORF in second and third grades suggest that individual differences in ORF growth rate in first grade strongly predict where students' ORF would begin in subsequent years. The critical variation in growth rate of ORF occurs during first grade, and in second and third grades the growth rate may remain relatively similar and stable across students, thus providing little variation in rank ordering.

The findings in this study are in line with the overall conceptual framework of the RTI approaches. According to the RTI framework, teachers screen and identify struggling readers based on limited response to effective instruction or intervention (Compton et al., 2006; D. Fuchs et al., 2003; Lyon et al., 2003). In addition to students' performance at the beginning of the year (i.e., fall), students' limited or low progress rate in ORF can facilitate identifying students in need of early intervention. This progress monitoring may be a particularly useful tool in first grade because students' ORF at the beginning of the year may have been depressed due to their limited word reading skills, although students' fall performance in first grade varied sufficiently to be related to their reading comprehension at the end of first and third grades. Despite these depressed ORF scores at the beginning of first grade, however, the fan-spread growth pattern in first-grade ORF (as observed by a positive correlation between initial status and growth rate in ORF) suggests that students who started at a low level of ORF in the beginning of first grade tend to have a slower growth rate across first grade. Thus, the number of words correctly read at the beginning of first grade may be a good indication of how fast students' ORF will grow during the first grade. These results offer two implications. First, struggling readers at this stage are likely to struggle with word decoding skills, and thus, based on further diagnostic assessments, these students should be provided with appropriate instruction. Second, teachers may need to pay close attention to students who struggle in the beginning of the year and monitor their rate of progress, although the WCPM difference between fluent and dysfluent readers might appear small in the beginning of the year.

Typically, teachers at Reading First schools were not required to assess students' ORF beyond three benchmark assessment periods

Table 7
Dominance Analysis Results From Second-Grade Prediction of Third-Grade SAT-10

Subset model	R^2	Unique contribution of predictor to third-grade SAT-10				
		ORFF	ORFS	ORFG	SAT-10	PPVT-III
Models with one predictor						
ORFF	.39		0.05	0.04	0.20	0.12
ORFS	.42	0.01		0.01	0.17	0.11
ORFG	.20	0.23	0.23		0.36	0.19
SAT-10	.53	0.05	0.06	0.02		0.04
PPVT-III	.31	0.19	0.23	0.08	0.26	
Models with two predictors						
ORFF–ORFS	.44			0.00	0.16	0.11
ORFF–ORFG	.42		0.02		0.17	0.11
ORFF–SAT-10	.59		0.01	0.01		0.03
ORFF–PPVT-III	.51		0.04	0.03	0.11	
ORFS–ORFG	.43	0.01			0.17	0.11
ORFS–SAT-10	.59	0.00		0.00		0.03
ORFS–PPVT-III	.54	0.01		0.00	0.09	
ORFG–SAT-10	.55	0.04	0.04			0.03
ORFG–PPVT-III	.39	0.15	0.15		0.20	
SAT-10–PPVT-III	.57	0.05	0.06	0.01		
Models with three predictors						
ORFF–ORFS–ORFG	.44				0.16	0.11
ORFF–ORFS–SAT-10	.60			0.00		0.03
ORFF–ORFS–PPVT-III	.54			0.00	0.09	
ORFF–ORFG–SAT-10	.60		0.00			0.03
ORFF–ORFG–PPVT-III	.54		0.01		0.09	
ORFF–SAT-10–PPVT-III	.62		0.01	0.01		
ORFS–ORFG–SAT-10	.60	0.00				0.03
ORFS–ORFG–PPVT-III	.54	0.01			0.09	
ORFS–SAT-10–PPVT-III	.63	0.00		0.00		
ORFG–SAT-10–PPVT-III	.59	0.04	0.04			
Models with four predictors						
ORFF–ORFS–ORFG–SAT-10	.60					0.03
ORFF–ORFS–ORFG–PPVT-III	.55				0.09	
ORFF–ORFS–SAT-10–PPVT-III	.63			0.00		
ORFF–ORFG–SAT-10–PPVT-III	.63		0.00			
ORFS–ORFG–SAT-10–PPVT-III	.63	0.00				
Model with five predictors						
ORFF–ORFS–ORFG–SAT-10–PPVT-III	.63					

Note. The letters *F*, *S*, and *G* after a variable name represent fall, spring, and growth, respectively. ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

(fall, winter, and spring, although there were four benchmark assessments for the data used in the present study). This frequency of assessment may suffice for most students who are on track in their reading development. However, the importance of growth rate in first grade suggests that it may be necessary to assess first-grade students more frequently, particularly those who are at risk based on a fall screening assessment (determined by low ORF and other assessments such as nonword reading fluency, LNF, and phonemic segmentation). A recent study suggests that progress monitoring data with ORF collected every 3 weeks provide estimates that are closest to true growth rates compared to measurements taken weekly, biweekly, and every 4 weeks and 9 weeks as long as multiple passages are used at baseline and other monitoring points (Jenkins, Graff, & Miglioretti, 2009; but see Shinn, 2002, for more frequent assessments). Given the brief nature of the ORF assessment, close monitoring of ORF either for the entire class or selected at-risk students may be feasible for educators (e.g., classroom teachers or reading coaches) without reallocating resources.

The ultimate purpose of a curriculum-based measurement ORF assessment is to facilitate enhanced instructional planning (Deno, 1985), and create "a database for each student to allow the teacher to evaluate the effectiveness of an individual student's educational program" (Deno, 1992, p. 5). Thus, results from screening and progress monitoring assessments should guide prevention and early intervention activities (Christ & Silberglitt, 2007) as well as serve as a primary data source for high-stakes decisions such as diagnostic and eligibility decisions (Speece, 2005). In keeping with an RTI framework, the manner and intensity of intervention or instruction is adjusted based on the student's rate of progress (see Vaughn, Wanzek, Woodruff, & Linan-Thompson, 2007, for more details). When students are provided with a research-based core curriculum (i.e., classroom or Tier 1 instruction) and the growth rate appears discrepant from what is typically achieved by other students, the instruction is adjusted by increasing the time and intensity of instruction in the core curriculum (i.e., Tier 2 instruction). For example, students can be provided with supplementary 30-min instruction on target areas every day in a small group. The teacher monitors these students' progress rate, which, then, informs further steps to be taken about instruction (e.g., more intense supplementary instruction such as Tier 3 one-on-one tutoring or return to Tier 1 instruction). Thus, frequent growth rate monitoring would allow educators to flexibly respond to students' changing instructional needs, whereas a single point assessment could conceal a potentially more dynamic pattern of students' learning.

In contrast to first grade, students' level of ORF in the beginning of second and third grades was the dominant predictor of their later reading comprehension achievement. This implies that by the beginning of second grade, ORF growth rates are stable so that students' level in the beginning of the year provides sufficient information for screening purposes to determine which students need intensive individualized interventions. It is likely, however, that children with the lowest initial fluency scores (i.e., the lowest 30% or those at high risk) who would likely be receiving Tier 2 or 3 interventions would still benefit from frequent monitoring in order to modify their intervention and to determine when and if they read well enough to return to Tier 1 instruction.

Several limitations in the present study should be noted. The present study did not take into account form effects of passages used in the DIBELS ORF measure. Recently it was shown that passages used in the DIBELS ORF assessment should be empirically equated (i.e., reading rate) for longitudinal estimation of ORF development (Francis, Santi, Barr, Fletcher, Varisco, & Foorman, 2008). However, the impact of passage effects on growth modeling is not clear yet and future studies with a large sample are needed to investigate the effects of the level of passage difficulty used for equating (i.e., easiest, median, and most difficult). In addition, the sample of students in the present study were from Reading First schools in Florida with a large number of students from low socioeconomic backgrounds, and they were taught by teachers who had received training in Florida's Reading First implementation and use of selected curricula. Thus, the generalizability of the findings is limited to students of similar backgrounds. Future studies should replicate the findings of the present study with equated passages and with students of diverse socioeconomic backgrounds.

In conclusion, the present study demonstrated that in addition to students' assessment results in the beginning of the year, divergence in students' ORF growth rate, particularly during the course of first grade, reveals important prognostic information about students' later reading comprehension achievement. The findings of the present study provide empirical evidence that ORF growth rate in first grade may have a potential to facilitate making instructional decisions.

Table 8
Dominance Analysis Results From Prediction of Third-Grade SAT-10

Subset model	R^2	Unique contribution of predictor to third-grade SAT-10		
		ORFF	ORFG	PPVT-III
Models with one predictor				
ORFF	.44		0.00	0.11
ORFG	.15	0.30		0.24
PPVT-III	.34	0.21	0.05	
Models with two predictors				
ORFF–ORFG	.44			0.11
ORFF–PPVT-III	.55		0.00	
ORFG–PPVT-III	.38	0.17		
Models with three predictors				
ORFF–ORFG–PPVT-III	.55			

Note. ORFF and ORFG represent ORF fall performance and growth, respectively. ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

Table 9
95% Confidence Intervals for Pairwise Differences

Variables compared	ΔR^2	*SE*	95% CI	
			Lower limit	Upper limit
First grade				
ORFG–ORFF	.17	0.003	0.164	0.176
All	.05	0.002	0.046	0.054
ORFG–ORFS	.06	0.001	0.058	0.062
All	.05	0.002	0.046	0.054
ORFG–SAT-10	.01	0.0002	0.009	0.011
All	.02	0.001	0.017	0.023
ORFG–PPVT-III	.15	0.003	0.145	0.156
All	.01	0.001	0.008	0.012
Second grade				
ORFG–ORFF	−.16	0.003	−0.165	−0.155
All	.00	0.001	−0.001	0.001
ORFG–ORFS	−.22	0.004	−0.227	−0.213
All	.00	0.001	−0.001	0.001
ORFG–SAT-10	−.34	0.005	−0.350	−0.330
All	−.08	0.003	0.074	0.089
ORFG–PPVT-III	−.11	0.002	−0.114	−0.106
All	−.03	0.002	0.027	0.033
Third grade				
ORFG–ORFF	−.30	0.005	−0.309	−0.291
All	−.17	0.004	−0.178	−0.162
ORFG–PPVT-III	−.19	0.003	−0.196	−0.184
All	−.11	0.003	−0.119	−0.106

Note. ORFF and ORFG represent ORF fall performance and growth, respectively. ORF = Oral Reading Fluency; PPVT-III = Peabody Picture Vocabulary Test—Third Edition; SAT-10 = Stanford Achievement Test (10th ed.).

References

Alf, E. F., Jr., & Graf, R. G. (1999). Asymptotic confidence limits for the difference between two squared multiple correlations: A simplified approach. *Psychological Methods, 4,* 70–75.

Buck, J., & Torgesen, J. (2003). The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test (FCRR Tech. Rep. No. 1). Tallahassee, FL: Florida Center for Reading Research.

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin, 114,* 542–551.

Christ, T. J., & Silberglitt, B. (2007). Estimates of standard error of measurement for curriculum-based measures of oral reading fluency. *School Psychology Review, 36,* 130–146.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98,* 394–409.

Cook, R. G. (2003). *The utility of DIBELS as a curriculum based measurement in relation to reading proficiency on high-stakes tests.* Unpublished master's thesis, Marshall University.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Deno, S. L. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure, 36,* 5–10.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test—Third Edition.* Bloomington, MN: Pearson Assessments.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40,* 532–538.

Francis, D. J., Santi, K. L., Barr, C., Fletcher, J. M., Varisco, A., & Foorman, B. R. (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology, 46,* 315–342.

Fuchs, D., Fuchs, L. S., McMaster, K. N., & Al Otaiba, S. (2003). Identifying children at risk for reading failure: Curriculum-based measurement and the dual-discrepancy approach. In H. L. Swanson & K. R. Harris (Eds.), *Handbook of learning disabilities* (pp. 431–449). New York, NY: Guilford Press.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice, 13,* 204–219.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral Reading Fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5,* 239–256.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement. Retrieved from http://dibels.uoregon.edu

Good, R. H., Kaminski, R. A., Smith, S., Laimon D., & Dill, S. (2001). *Dynamic Indicators of Basic Early Literacy Skills* (5th ed.). Eugene, OR: University of Oregon.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency based indicators of foundational reading skills for third-grade high stakes outcomes. *Scientific Studies of Reading, 5,* 257–288.

Harcourt Brace. (2003). *Stanford Achievement Test—Tenth Edition: Technical data report.* San Antonio, TX: Author.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95,* 719–729.

Jenkins, J. R., Graff, J. J., & Miglioretti, D. L. (2009). Estimating reading growth using intermittent CBM progress monitoring. *Exceptional Children, 75,* 151–163.

Jenkins, J. R., & Jewell, M. (1993). Examining the validity two measures for formative teaching: Reading aloud and maze. *Exceptional Children, 59,* 421–432.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25,* 215–227.

Kaminski, R. A., & Good, R. H. (1998). Use of curriculum-based measurement to assess early literacy: Dynamic indicators of basic early literacy skills. In M. Shinn (Ed.), *Advances in curriculum-based measurement and its use in a problem-solving model.* New York, NY: Guilford Press.

Katzir, T., Shaul, S., Breznitz, Z., & Wolf, M. (2004). The universal and unique in dyslexia: A cross-linguistic investigation of reading and reading fluency in Hebrew- and English-speaking children with reading disorders. *Reading and Writing: An Interdisciplinary Journal, 17,* 739–768.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53,* 1–14.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71,* 445–463.

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

Neter, J., Kutner, M. H., Wasserman, W., & Nachtsheim, C. J. (1996). *Applied linear statistical models* (4th ed.). New York, NY: McGraw-Hill.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Perfetti, C. A. (1985). *Reading ability.* New York, NY: Oxford University Press.

Raudenbush, S. W., & Bryk, A. S. (2001). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). HLM 6.06 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International.

Ridel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly, 42,* 546–567.

Roberts, G., Good, R., & Corcoran, S. (2005). Story retell: A fluency-based indicator of reading comprehension. *School Psychology Quarterly, 20,* 304–317.

Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Not just speed reading: Accuracy of the DIBELS oral reading fluency measure for predicting high-stakes third grade reading comprehension outcomes. *Journal of School Psychology, 46,* 343–366.

Schatschneider, C., Wagner, R. K., & Crawford, E. C. (2008). The importance of measuring growth in response to intervention models: Testing a core assumption. *Learning and Individual Differences, 18,* 308–315.

Shinn, M. R. (2002). Best practices in curriculum-based measurement. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 671–698). Bethesda, MD: National Association of School Psychologists.

Speece, D. L. (2005). Hitting the moving target known as reading development: Some thoughts on screening children for secondary interventions. *Journal of Learning Disabilities, 38,* 487–493.

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93,* 735–749.

Speece, D. L., Mills, C., Ritchey, K. D., & Hillman, E. (2003). Initial evidence that letter fluency tasks are valid indicators of early reading skill. *Journal of Special Education, 36,* 223–233.

Stahl, S. A., & Kuhn, M. R. (2002). Making it sound like language: Developing fluency. *The Reading Teacher, 55,* 582–584.

Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial. *Educational & Psychological Measurement, 55,* 525–534.

Vaughn, S., Wanzek, J., Woodruff, A. L., & Linan-Thompson, S. (2007). Prevention and early identification of students with reading disabilities. In D. H. Haager, S. Vaughn, & J. K. Klinger (Eds.), *Evidence-based reading practices for response to intervention* (pp. 11–27). Baltimore, MD: Brookes.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848–872.

Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26,* 207–214.

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

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,* 4–12.

Appendix A

Level 1 Model

$$\text{SAT-10}_{ti} = \pi_{0i} + \pi_{1i}(\text{Month}_{ti}) + \pi_{2i}(\text{Month}^2_{ti})$$

$$+ \pi_{3i}(2^{nd}\text{ Grade Int}_{ti}) + \pi_{4i}(3^{rd}\text{ Grade Int}_{ti})$$

$$+ \pi_{5i}(2^{nd}\text{ Grade Lin}_{ti}) + \pi_{6i}(3^{rd}\text{ Grade Lin}_{ti})$$

$$+ \pi_{7i}(2^{nd}\text{ Grade Quad}_{ti}) + \pi_{8i}(3^{rd}\text{ Grade Quad}_{ti}) + e_{ti}$$

Level 2 Model

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$\pi_{2i} = \beta_{20} + r_{2i}$$

$$\pi_{3i} = \beta_{30} + r_{3i}$$

$$\pi_{4i} = \beta_{40} + r_{4i}$$

$$\pi_{5i} = \beta_{50} + r_{5i}$$

$$\pi_{6i} = \beta_{60} + r_{6i}$$

$$\pi_{7i} = \beta_{70} + r_{7i}$$

$$\pi_{8i} = \beta_{80} + r_{8i}$$

Note. SAT-10 = Stanford Achievement Test (10th ed.); Int = intercept; Lin = linear term; Quad = quadratic term.

Appendix B

Monthly Growth Rates in NWF in First and Second Grades and ORF in First, Second, and Third Grades

	ORF			NWF	
Month	First grade	Second grade	Third grade	First grade	Second grade
September	2.99	4.40	5.63	4.81	5.28
October	3.50	4.71	5.63	4.72	4.98
November	4.00	5.01	5.63	4.64	4.68
December	4.51	5.31	5.63	4.55	4.38
January	5.02	5.62	5.63	4.47	4.08
February	5.53	5.92	5.63	4.38	3.78
March	6.03	6.23	5.63	4.29	3.48

Note. Rate of celeration for ORF in first grade = 0.25; rate of celeration for NWF in first grade = −0.04; rate of celeration for ORF in second grade = 0.15; rate of celeration for NWF second grade = −0.15; rate of celeration for ORF in third grade = 0.00. ORF = Oral Reading Fluency; NWF = Nonsense Word Fluency.

Received November 4, 2008
Revision received March 18, 2010
Accepted March 29, 2010 ■