RUNNING HEAD: Progress Monitoring with CBM

Within-Year Achievement Growth Trajectories Using Progress Monitoring Measures

Gerald Tindal
Joseph F. T. Nese
Joe Stevens
Julie Alonzo

University of Oregon

Acknowledgments:
Note: This project was funded through the National Center on Assessment and Accountability for Special Education (NCAASE) grant (R324C110004) from the U.S. Department of Education, Institute of Education Sciences. Opinions expressed do not necessarily reflect the opinions or policies of the U.S. Department of Education.

Abstract

For 30 years, researchers have investigated various properties of oral reading fluency documenting content- and criterion-related evidence supporting its use as a measure of performance. Another line of research has addressed its adequacy as a measure of growth. Although the average weekly growth has generally been consistent, little research has been done with formative measures used in school contexts and in relation to either measurement conditions or student characteristics. We use a two-level hierarchical linear model to document at level-1 the effects of time and at level-2, the effects of measurement conditions (grade level of the measure used to monitor progress and the number of progress measures administered) and student characteristics (sex, special education services, race-ethnicity, and English language learner status). For students in three grades, we found considerable consistency in the effects from the number of assessments administered and for students receiving special education services.

The Influence of Student Characteristics and Measurement Conditions on Growth in Oral Reading Fluency: Progress Monitoring with CBM

This paper is a follow up to the paper presented by Tindal and Nese (2012) (Tindal & Nese)(Tindal & Nese)(Tindal & Nese)(Tindal & Nese)(Tindal & Nese)(Tindal & Nese)in which the same type of analysis was conducted with easyCBM data from Grade 4 students. Three changes are reflected in the current paper.

1. The sampling plan in the current study is more restrictive in selecting only students with the first progress measurement in September and at least seven progress measures administered throughout the year in the same grade level (either on grade or one grade below). Therefore, the results differ slightly for fourth grade students. In addition, students in Grades 3 and 5 also were sampled using this same set of requirements.
2. The literature sampled on oral reading fluency (ORF) is considerably expanded beyond the eight studies included in the previous paper, with seven more studies summarized in the table. In addition, literature on relevant student characteristics is added to this paper, including the over representation of students of color, males, and English language learners in special education.
3. We more clearly describe models for testing growth in this paper but do not report a non-linear model as the quadratic model resulted in non-significant *p*-values for slope in the final estimation of fixed effects in grades 3 and 4.

The Early Research on Oral Reading Fluency Growth

In a number of early studies (1980s and 1990s), growth in oral reading fluency was typically documented as pre-post differences obtained under varying (treatment) conditions. For example, in an experimental study conducted from November to May, = two groups of teachers were compared: (a) systematic IEP goal monitoring (changing ineffective programs every 7-10 measurement points) and (b) IEP monitoring as teachers wished (L. Fuchs, Deno, & Mirkin, 1984). At the end of the study a pre-post comparison was made on fluency with 1 word correct per minute (wcpm) increase over 28 weeks obtained with systematic IEP monitoring. In another experimental study with students assigned to one of three conditions (explicit teaching, explicit teaching and peer tutoring, and a control condition); improvement over 15 weeks was 2.6 wcpm, 4.4 wcpm, and 2.4 wcpm, respectively (Simmons, Fuchs, Fuchs, Mathes, & Pate, 1995).(Simmons et al., 1995) In an experimental study (comparing peer assisted learning strategies [PALS] with no PALS), D. Fuchs, Fuchs, Mathes, and Simmons (1997a) documented reading fluency changes over 15 weeks for students with learning disabilities compared with low performing and average achieving students. They reported that with PALS, these groups showed improvements of 51 wcpm, 47 wcpm, and 60 wcpm, respectively. These rates of improvement were considerably less without PALS (28 wcpm, 40 wcpm, and 37 wcpm, repectively). In another study of PALS, the gains in oral reading fluency were considerably less per week, averaging .56 wcpm (Fuchs, Fuchs, Mathes, & Simmons, 1987), which is considerably lower than previously cited expectations of 2.10 wcpm (Fuchs, Fuchs, Hamlett, & Ferguson, 1992).

In all later, studies, slope is calculated not as pre-post differences but as a weekly slope of improvement obtained by multiplying calendar days by seven, thus providing average weekly increase across the school year. In the first study to document typical growth, “seven generic passages were used” (p. 31) to answer the question of ‘how much growth can we expect’ (Fuchs, et al., 1993)? Results were analyzed by grade level using ordinary least squares (OLS) regression. The lowest amount of growth was in Grade 6 (with .32 wcpm per week) and the greatest amount was in Grade 1 (with 2.10 wcpm per week). Even though the sample sizes were small in each grade (ranging from 16 to 25), the idea was important and set the stage for others to follow (Fuchs, et al., 1993).

In a similar study eight years after the first one, Deno et al. (2001) reported growth on oral reading fluency in an effort to establish growth standards. They provided slopes at various levels of intercept for both general and special education students. Average growth was consistently higher at the lower levels of intercept and consistently different between general and special education students. While general education students had on average above 1.0 wcpm (and up to 1.91 in Grade 1), special education students were half that with the highest growth .76 wcpm in Grade 1.

Recent Research on Oral Reading Fluency Growth

Given the ubiquity of ORF as a screener, it is important to use these data for defining expectations on growth throughout the year. In a study by Ardoin and Christ (2008), DIBELS passages were adminstered to second grade students. They correlated performance across the three time periods (fall, winter, and spring), and using difference scores (between fall and winter versus winter and spring) conducted significance tests to determine if growth was consistent. They reported more growth in the first trimester than the second trimester. For this group of second grade students, weekly growth was consistently above 1 wcpm per week from fall to winter and less than 1 wcpm per week from winter to spring; these differences were significant. In a slight refinement of methodology, Christ, Silberglitt, Yeo, and Cormier (2010) studied 'seasonal' effects (fall-winter-spring benchmarks) for students in Grades 2-6 in both general and special education. Rather than difference scores, they used both linear and piecewise models to ascertain whether this growth was consistent. They reported that the "piecewise model fit significantly better than the linear model" (p. 453) with much less growth occurring in the second semester. For Graney, Missall, Martinez, & Bergstrom (2009), the focus was on growth for benchmark measures using AIMSweb passages (fall-winter-spring) for both general and special education students in Grades 3-5. Their intervals were unequal so corrections were made to make the semesters comparable. They used a "3 (grade level) by 2 (time) repeated measures ANOVA [and] demonstrated a significant interaction between grade level and time in Year 1, $F(2, 439)=3.72$, $p<.05$. Growth rates for third grade students increased negligibly from the first to the second half of the year, whereas the fourth and fifth grade students' growth increased substantially" (p. 132). No such significant interaction appeared in Year 2, although a main effect for benchmark periods was significant. They also reported "students in general education demonstrating higher rates of growth than their counterparts in special education" (p. 135). Finally, Nese, Biancarosa, Anderson, Lai, Alonzo and Tindal (2012) reported nonlinear slopes (with more growth from fall to winter than from winter to spring) but also more growth in the early grades over later grades.

Three studies have been published that have investigated growth in ORF in a very proximal manner (over weeks rather than the course of the year) and in relation to other variables. Although the purpose of the research by Jenkins, Zumeta, and Dupree (2005) was on passage comparability, they raised a critical issue when interpreting growth: "Participants demonstrated

PRF gains of just over 1 wcpm per week, whether calculated for the 5- or 10-week testing intervals. Gains did not differ significantly between same and different passages for either interval (p. 251). An important caveat, however, was that this gain was only about twice that of the standard error (6-8 wcpm). The study by Jenkins and Terjesen (2011) addressed goal ambitiousness, monitoring frequency, and method of evaluation. Most importantly for purposes of our current study, "across 8 weeks of monitoring, the sample registered a mean growth slope between 1.48 and 1.67 WRC per week, depending on whether only baseline and week 8 scores or all scores were used in the slope calculation" (p. 33). They also reported significant effects from the use of ambitious goals (in the number of changes to instruction) and an interaction with frequency of monitoring and evaluation method (with slope generating more changes in instruction than data points below the slope). Finally, Jenkins, Graff, and Miglioretti (2009) "compared slopes generated from five distinct progress monitoring schedules: one passage every week (weekly); two passages every 2 weeks; three passages every 3 weeks; four passages every 4 weeks; and first/last weeks only" (p. 155). Using the 'true' slope obtained from the full complement of 29 passages, they reported mean growth of 1.09 wcpm. The correlations of all monitoring schemes were generally high when more than one baseline and more than one monitoring probe, and all of the five monitoring schemes were comparable to the true slope. Interestingly, every 3 weeks and first-last week probes most closely approximated slope.

Documenting student characteristics in oral reading fluency growth. Most of the research on oral reading fluency growth has not considered student demographics in their analysis (Ardoin & Christ, 2008; L. Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Jenkins et al., 2009; Jenkins & Terjeson, 2011; Jenkins et al., 2005). Three studies have been completed in which students receiving special education services have been compared with those in general education (Christ, 2006; Deno, Fuchs, Marston, & Shin, 2001; Graney, Missall, Martinez, & Bergstrom, 2009). The only study to analyze performance and progress with respect to student characteristics was published by Nese et al. (2012) who reported differences at the beginning of the year with higher performance for female students (4 more wcpm than males), lower performance for students eligible for free and reduced-price lunch (11 wcpm less), lower performance for students with diabilities (37 wcpm less), and lower performance for limited English proficient students (20 wcpm less). Slope was not different among these groups except for students with disabilities who improved from fall to winter by about 4 wcpm less than general education students.

Yet, it is commonly documented that special education is rife with problems of over-representation of various groups of students, particularly *students of color*. For example, "the disproportionate representation of minority students is among the most critical and enduring problems in the field of special education" (Skiba et al., 2008, p. 264). Beginning with *Brown v. Board of Education (1954),* continuing with *Larry P. v. Riles (1972),* and now most recently with the Office of Civil Rights (OCR) from the 1980s to the present, disproportionality has been highlighted as problematic. Indeed, the problem is so significant that the reauthorization of IDEA (2004) requires states to "monitor disproportionate representation by race or ethnicity in disability categories and special education placements and require the review of local policies, practices, and procedures when disproportionate representation is found" (p. 266).

Typically, African American students are more than twice as likely to be identified with mental retardation than are actually present in the population and this finding has been consistent over time; furthermore, this overrepresentation tends to occur more in restrictive than in non-restrictive settings (Skiba et al., 2009). According to (Vallas, 2009) Vallas (2009), African Americans represent 15% of the student-age population but 20% of students with disabilities. In the most recent (30th) *Annual Report to Congress* (Office of Special Education and Rehabilitative Services, 2011), the largest group of students (ages 6 through 21) identified with Intellectual Disabilities were Black (nearly 14%) while all other races were around 7% (*N.B.* These data were collected in Fall 2006). "In 2006, the percentage of the population ages 6 through 21 served under *IDEA*, *Part B,* varied by race/ethnicity. The percentage served (i.e., risk index) was largest for American Indian/Alaska Native students (13.86 percent), followed by Black (not Hispanic) students (12.22 percent), White (not Hispanic) students (8.52 percent), Hispanic students (8.41 percent) and Asian/Pacific Islander students (4.66 percent)" (p. 52). Furthermore, two subgroups had quite high-risk ratios (likelihood of being served under IDEA, Part B): American Indian/Alaska Native students were 1.56 times more likely and Black (not Hispanic) students were 1.46 times more likely to be served in general; for intellectual disabilities, Blacks were 2.75 times more likely to be served and for emotional disturbance, 2.28 times more likely to be served. Clearly, the issue of race and identification with special needs still tends to reflect systematic variance. Generally, this outcome is viewed negatively but as Feldman (2011) Feldman (2011) notes "disproportionality alone is not the only problem but may be one of other problems that involve the cause as well as the consequences of such overrepresentation" (p. 186). For example, causes may be psychometric (e.g., item writing), socio-demographic (e.g. poverty), or opportunity to learn (Skiba et al., 2008).

Over-representation of *males* in special education also is prevalent and has been for many years. For example, Zorigian and Job (Oct. 12, 2012) reported that in most disability categories, males are identified with a disability more than girls; in most categories, they are 2-4 times as likely with the exception of sensory impairments where the rates are only slightly higher. In 2003, males were more than twice as likely to be labeled with learning disabilities, more than three times as likely to be identified with emotional disturbance, and more than seven times as likely to be labeled with autism.

Finally, over-representation of *English language learners* in special education is present (sometimes conflating students in these two populations). Artiles, Rueda, Salazar, and Higareda (2005), note that although "the proportion of ELLs receiving special education services in the [urban] target districts was comparable to the total student populations (7.6%) a much greater proportion of secondary students received special education service (14.1%) than elementary students (5.3%)" (p. 287). They reported risk index data that was highest for white students with 11.2% placed in special education, followed by two language-based groups: 7.2% of English language learners and 6.8% English proficient learners. Again, they noted that the risk index was highest for ELLS in secondary grades. Essentially, ELLs are under-represented in special education grades in elementary schools and over-represented in secondary schools.

Summary. Most of the research on growth of oral reading fluency has tended to focus on benchmark tests within the year with measures given three times: In the fall, winter, and spring. The only exception is the research by Jenkins and colleagues who measured more frequently but within an overall shorter time period. None of the previous studies have drawn from progress monitoring data as operationalized by teachers in the field (using a convenience sample but in a naturalistic environment). Furthermore, the research has not been detailed in the analyses with respect to student characteristics. Though samples have been described in terms of student characteristics, they have not been incorporated into the analyses. Therefore, the purpose of this study is to take a sample of teachers who are progress monitoring students as part of their response-to-intervention and document the influence of time and student characteristics on intercept (beginning performance in the fall) and slope (change over time throughout the year as teachers monitor progress). See Table 1 for a summary of research on oral reading fluency growth.

Methods

Population and students sampled. Participants for the study included Grade 3-5 students in all districts using easyCBM at the end of the 2010-2011 school year. easyCBM is an assessment system that provides teachers and administrators performance and progress information on basic skills in reading and mathematics. In all, easyCBM district users in 2010-2011 included 24 grade three, 19 in grade four, and 21 in grade five.

In grade three, a total of 2,332 students were in the system taking progress measures; many of these students also had a performance assessment (fall, winter, or spring benchmark). Progress measure scores were deleted for five students as out of range. We removed -6 in score 1, -108 in score 3, -84 in score 3, 11107 in score 7, and 741 in score 10.

In grade four, 2,601 students were in the system taking progress measures; again, many of these students also had a performance assessment (fall, winter, or spring benchmark). One progress measure score was deleted as out of range. We removed 853 in score 5.

Finally, in grade five, there were 1,909 students taking progress measures. As in grades three and four, many of these students also had a performance assessment (fall, winter, or spring benchmark). One progress measure score was deleted as out of range. We removed -447 in score 6.

Dependent measure The measures used in this study consisted of only passage reading fluency (PRF).[1] The easyCBM PRF measures used in this study have an extensive number of technical reports available at http://brtprojects.org. When districts enroll in the system, a student roster is uploaded. Generally, districts require all students to take benchmark measures in the fall, winter, and spring, and then teachers use this system to select students at risk of failure. Teachers have the capacity to assign students into (reading) groups and select progress monitoring measures appropriate for the student's skill level. They are provided measures to download and administer in a standardized manner with the easyCBM web site providing training examples for administration and scoring of all reading measures. Once students are measured, teachers enter the data into the computer and the values are plotted in a time series graph. Teachers also have the option of introducing instructional programs into this time series by providing a label (that appears on the graph with a vertical line on the day of its introduction) and a description of it (accessible through a related report).

[1] Passage reading fluency (PRF) here is used as a synonym for oral reading fluency (ORF) to differentiate it from word reading fluency. Both PRF and ORF describe an individually administered test of accuracy and fluency with connected text.

Data analysis. We used HLM version 6.08 to estimate all models. We first examined unconditional models (without predictors but with time metric variable(s)) to examine mean and variance of within-subject reading fluency and to provide baseline statistics for evaluating subsequent conditional models (Raudenbush & Bryk, 2002). The two parameters of interest in level-1 represented intercept and slope, with slope centered on the fall performance in the grade level (first progress measure in September) and slope expressed as the weekly increase over the year. We first examined fixed effects and then random effects.

We ran conditional models at level-2 to determine the influence of four student characteristics and two measurement conditions on both the intercept and slope. Student characteristics were coded as: (a) student sex, with female = 0, male =1; (b) program placement, with general education=0 and special education = 1; (c) student ethnicity-race, with White=0 and Non-White =1; and (d) English language learner status, with English speaker = 0 and English language learner = 1. Two measurement conditions were coded as: (a) grade level of progress monitoring measures (0=on-grade and 1= lower grade) and (b) number of passage reading fluency assessments administered (a count from 1-33). After documenting descriptive statistics for the measures and the student characteristics, we ran a two-level hierarchical linear growth model.

Unconditional Models:

Level 1 $Y_{ti} = \pi_{0i} + \pi_{1i}(\text{time}) + e_{ti}$

Level 2 $\pi_{0i} = \beta_{00} + r_{0i}$

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

Conditional Models:

Level 1 $Y_{ti} = \pi_{0i} + \pi_{1i}(\text{time}) + e_{ti}$

Level 2 $\pi_{0i} = \beta_{00} + \beta_{01}(\text{Sex}) + \beta_{02}(\text{Disability}) + \beta_{03}(\text{Ethnicity}) + \beta_{04}(\text{GrLevel}) + \beta_{05}(\text{PRFs}) + r_{0i}$

$\pi_{1ib} = \beta_{10} + \beta_{11}(\text{Sex}) + \beta_{12}(\text{Disability}) + \beta_{13}(\text{Ethnicity}) + \beta_{14}(\text{GrLevel}) + \beta_{15}(\text{PRFs}) + r_{1i}$

where:

Y_{ti} is the outcome (i.e., assessment score) at time t for student i
π_{0i} is the status of student i at the first measurement occasion
π_{1i} is the linear growth rate over time for student i
β_{01} to β_{05} are the coefficients for student characteristics and measurement conditions for intercept
β_{11} to β_{15} are the coefficients for student characteristics and measurement conditions for slope
e_{ti} is a residual term representing unexplained student variation from the latent growth trajectory

Results

Descriptive statistics are reported in Tables 2-4 for intercept and slope used in the unconditional model and all of the measurement conditions and student characteristics used in the conditional models. In all three grades, time was analyzed with both linear and quadratic functions in the initial analysis and then only linear function considered as the baseline before adding in student and measurement characteristics. In Grades 3 and 4, when time was considered nonlinear, no significant effects were found; in grade five, though significant, the linear function was used as baseline to make the results comparable across all elementary grades.

Grade 3. Results show that the average oral reading fluency across all Grade 3 students was 72.03. On average, for each additional week, fluency increased .67 wcpm. Deviance was 18692.53 with six estimated parameters. See Table 5.

When student and measurement characteristics were included, the average ORF across all students was 99.583words correct per minute (wcpm) ($SE = 3.11$). Only one student characteristic was significantly related to intercept: Disability, $t(607) = -5.03$, $SE = 4.50$, $p < .001$. Controlling for all other student characteristics, students with disabilities read nearly 23 fewer wcpm at fall. In contrast, both measurement conditions were significant predictors of intercept, controlling for all student characteristics: (a) grade level measured, $t(607)=-3.19$, $SE=4.33$, $p=.002$, and (b) number of passage reading fluency assessments, $t(607)=-10.47$, $SE=.38$, $p<.001$. Students measured below grade level read on average 14 fewer wcpm than students on grade level. In the fall, students read on average nearly four fewer wcpm with each additional weekly progress measure.

The overall slope was .21 wcpm growth per week, controlling for all student characteristics and both measurement conditions; this amount of growth was significant with $t(607) = 2.16$, $SE =.100$, $p =.031$. Otherwise, the only characteristic (student or measurement) with significant influence on slope was the number of passage reading fluency assessments: $t(607) = 7.28$, $SE = .01$, $p < .001$. For each additional week of assessment administration, students increased .07 wcpm. As random effects, the variances of intercept and slope were significant in the final estimation of the conditional model. Deviance was 18462.45 with 18 estimated parameters; this model was significantly different than the baseline model ($X^2 = 230.08$, 12 $df = 12$). See Table 6.

Grade 4. Results show that the average oral reading fluency across all Grade 4 students was 97.10 wcpm. On average, for each additional week, fluency increased .65 wcpm. Students varied significantly in the estimates of fall intercept and slope. Deviance was 24751.13 with six estimated parameters. See Table 7.

Results show that the average fall ORF across all Grade 4 students was 115.61 wcpm with $t(952) = 42.83$, $SE = 2.70$, $p < .001$), controlling for student characteristics and both measurement conditions. Two student characteristics were significantly related to intercept: (a) Disability, $t(952) = -6.82$, $SE = 3.71$, $p < .001$, and (b) English language learner, $t(952) = -6.11$, $SE = 3.90$, $p < .001$. Students with disabilities averaged 25 fewer wcpm and those labeled ELL read nearly 24 fewer wcpm than peers when controlling for all other student and measurement conditions. Both measurement conditions were significantly related to intercept, controlling for all student characteristics: (a) grade level measured ($t(952) = -2.67$, SE=4.13, $p = .008$, and (b) number of number of passage reading fluency assessments, $t(952) = -6.60$, SE =.38, $p < .001$. Students read 11 fewer wcpm when measured below grade level and 2.5 fewer wcpm for every weekly assessment administered.

Controlling for both student characteristics and measurement conditions the average Grade 4 slope was .72 wcpm ($t(952) = 8.87$, $SE = .08$, $p < .001$).The only student characteristic that showed a significant effect on slope of improvement over time was ELL: ELL students averaged 1 wcpm increase every 4 weeks ($t(952) = 1.98$, $SE = .13$, $p = .047$). The random effect of intercept but not slope was significant in the final estimation of variance components for the conditional model. Deviance was 24474.09 with 18 estimated parameters; this model was significantly different than the baseline model ($X^2 = 277.04$, $df = 12$). See Table 8.

Grade 5. Results show that the average ORF across all Grade 5 students was 103.8 wcpm On average, for each additional week, fluency increased .79 wcpm. Deviance was 18147.32 with six estimated parameters. See Table 9.

Controlling for both student conditions and measurement conditions, the average ORF was 121.90 wcpm ($t(449) = 38.33$, $SE = 3.18$, $p < .001$). Only one student characteristic significantly predicted intercept: students with disabilities read on average fewer than 28 wcpm with $t(449) = -8.80$, $SE = 3.22$, $p < .001$. Both measurement conditions significantly predicted intercept controlling for all student characteristics: (a) grade level measured in which students measured below grade level read on average read nearly 18 fewer wcpm with $t(449) = -5.02$, $SE = 3.53$, $p <.001$ and (b) number of reading assessments in which the intercept was nearly 1 fewer wcpm with $t(449) = -2.12$, $SE = .42$, $p = .034$.

Controlling for all student characteristics and both measurement conditions, the slope was .96 wcpm per week ($t(449) = 12.31$, $SE = .08$, $p < .001$),. Only the number of passage reading fluency assessments administered was a statistically significant predictor of slope with $t(449) = -2.30$, $SE = .08$, $p = .022$). For each additional week of test administration, students read on average slightly fewer words correct (-.02 wcpm). Students varied significantly in the estimates of both intercept and slope. Deviance was 17942.18 with 18 estimated parameters; this model was significantly different than the baseline model ($X^2 = 205.13$, $df = 12$). See Table 10.

Discussion

The results from our analyses of students' ORF trajectories in Grades 3-5 showed that students with disabilities entered much lower at the beginning of the year but improved in a comparable manner in all grade levels. No other student characteristic reflected such a uniform effect, although initial performance (lower) and rate of improvement (higher) were significantly different from average for ELL students in grade 4. These results are somewhat discrepant from the research previously reported, with more consistency in the entry level (intercept) than in growth, but that is not surprising, give the sample here were students targeted for progress monitoring, and the samples for previous research tended to be all students taking universal screening measures.

For measurement conditions, students measured below grade level were significantly lower than those measured on grade level initially (controlling for all other student characteristics and number of progress assessments) but improved at a comparable rate over the year. The findings for the number of progress assessments were mixed. For some students (in 3rd and 5th grades) the differences were significant in the initial level of performance and in the slope of improvement; however, in Grade 4, only the initial performance was significantly lower than the intercept.

The only anomaly in our findings was the low slope for students in Grade 3, a value that is clearly discrepant from previous research, but again, the samples are arguably discrepant. Otherwise, the most remarkable finding in the literature on oral reading fluency is the similarity of growth over the year. From the initial findings of Fuchs, Deno, and Mirkin (1984) (1 wcpm for students in treatment) over 28 weeks to the most recent research by Nese, et al.(2012) with about 1 wcpm improvement, the literature reflects a sturdy rate of change within the school year. Of course, this general statement overlooks some of the nuances being reported such as seasonal effects (reflecting a non-linear trend) and the influence of student characteristics (Keller-Margulis, Clemens, Im, Kwok, & Booth, 2012; Nese et al., 2012). Nevertheless, given the great variation in the passages being used to measure oral reading fluency, the number of measurement occasions (each of which provides feedback to teachers), the populations being sampled (both in terms of grade level and personal characteristics), and the sophistication of the analyses (from per-post difference scores to piece-wise hiearchical models), the general finding of about 1 wcpm t over two weeks (.5 per week) to about 1.5 wcpm per week, appears to be the general outcome for students in grades 3-5.

In general, several student characteristics have a negative effect on ORF intercept and slope. Clearly, the most significant and consistent negative effect is the presence of a disability or the need for special services. This finding has been uniform over the 30 years of research (Deno et al., 2001; Kim, Petscher, Schatschneider, & Foorman, 2010; Nese et al., 2012; Puranik, Petscher, Al Otaiba, Catts, & Lonigan, 2008; Wang, Algozzine, Ma, & Porfeli, 2011). Additionally, students who are English language learners often show both lower levels of performance initially and grow at a lower rate (Al Otaiba et al., 2009). Poverty and risk also appear to be important student characteristics that influence both level of performance and rate of progress (Crowe, Connor, & Petscher, 2009; Logan & Petscher, 2010). It appears possible to mitigate this general outcome with frequent progress measures used to provide teachers goals and feedback on their instruction (Jenkins & Terjeson, 2011) or to provide specialized programs (D. Fuchs, Fuchs, Mathes, & Simmons, 1997b; Simmons et al., 1995). The findings reported in this current study, however, showed mixed results with progress monitoring: In Grade 3, a positive affect was found on slope of improvement as a function of the number of progress assessments, with improvement non significant in grade 4, and improvement significantly lower in grade 5.

Of course, our findings and virtually all other studies on growth ignore consideration of instruction. One of the few studies to address learning in the context of teaching is the investigation conducted by (Crowe et al., 2009). They reported significant curriculum differences among Reading Mastery, Scott Foresman, Harcourt, Success for All, Open Court, and Houghton Mifflin, which also interacted with students from different SES groups.

References

Al Otaiba, S., Petscher, Y., Pappamihiel, N. E., Williams, R., Dyrlund, A. K., & Connor, C. (2009). Modeling oral reading fluency development in Latino dtudents: A longitudinal study across second and third grade. *Journal of Educational Psychology, 101*(2), 315-329.

Ardoin, S., & Christ, T. (2008). Evaluating curriculum-based measurement slope estimates using data from triannual universal screenings. *School Psychology Review, 37*, 109-125.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children, 71*(3), 283-300.

Christ, T. (2006). Short-term estimates of growth using curriculum-based measurement of oral reading fluency: Estimating standard error of the slope to construct confidence intervals. *School Psychology Review, 35*(1), 128-133.

Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology, 47*, 187-214.

Deno, S., Fuchs, L., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*(4), 507-524.

Feldman, J. (2011). Racial Perspectives on Eligibility for Special Education for Students of Color Who are Struggling, is Special Education a Potential Evil or a Potential Good? *Journal of Gender, Social Policy & the Law, 20*(1), 183-200.

Fuchs, D., Fuchs, L., Mathes, P., & Simmons, D. (1997a). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34*(1), 174-206. doi: DOI: 10.3102/00028312034001174

Fuchs, D., Fuchs, L., Mathes, P., & Simmons, D. (1997b). Peer-assisted learning strategies: Making classrooms more responsive to diversity. . *American Educational Research Journal, 34*(1), 174-206. doi: 10.3102/00028312034001174

Fuchs, L., Deno, S., & Mirkin, P. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal, 21*, 449-460. doi: 10.3102/00028312021002449

Fuchs, L., Fuchs, D., Hamlett, C., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Reveiw, 22*, 27-48.

Graney, S., Missall, K., Martinez, R., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School Psychology, 47*, 121-142. doi: 10.1016/j.jsp.2008.12.001

Jenkins, J., Graff, J. J., & Miglioretti, D. (2009). Estimating reading growth using intermittent CBM progress monitoring. *Exceptional Children, 75*(2), 151-163.

Jenkins, J., & Terjeson, K. (2011). Monitoring reading growth: Goal setting, measurement frequency, and methods of evaluation. *Learning Disabiities Research & Practice, 26*(1), 28-35.

Jenkins, J., Zumeta, R., & Dupree, O. (2005). Measuring gains in reading ability with passage reading fluency. *Learning Disabilities Research & Practice, 20*(4), 245-253. doi: 10.1111/j.1540-5826.2005.00140.x

Keller-Margulis, M., Clemens, N. H., Im, M. H., Kwok, O., & Booth, C. (2012). Curriculum-based measurement yearly growth rates: An examination of English Language Learners and native English speakers. *Learning and Individual Differences, 22*, 799-805.

Kim, Y., Petscher, Y., Schatschneider, C., & Foorman, B. (2010). Does growth rate in oral reading fluency matting in predicting reading comprehension achievement? *Journal of Educational Psychology, 1002*(3), 652-667.

Logan, J. A. R., & Petscher, Y. (2010). School profiles of at-risk student concentration: Differential growth in oral reading fluency. *Journal of School Psychology, 48*(2), 163-186.

Nese, J. F. T., Biancarosa, G., Anderson, D., Lai, C. F., Alonzo, J., & Tindal, G. (2012). Within-year oral reading fluency with CBM: A comparison of models. *Reading and Writing, 25*, 887-915. doi: 10.1007/s11145-011-9304-0

Office of Special Education and Rehabilitative Services, U. S. D. o. E. (2011). 30th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, 2008. Washington, DC: Office of Special Education and Rehabilitative Services.

Puranik, C. S., Petscher, Y., Al Otaiba, S., Catts, H. W., & Lonigan, C. J. (2008). Development of Oral Reading Fluency in Children With Speech or Language Impairments : A Growth Curve Analysis. *Journal of Learning Disabilities, 41*, 545-560.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*. Thousand Oaks, CA: Sage.

Simmons, D., Fuchs, L., Fuchs, D., Mathes, P., & Pate, J. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning-disabled and low-performing students in regular classrooms. *The Elementary School Journal,, 95*(5), 387-408.

Skiba, R. J., B., S. A., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., & Chung, C. G. (2008). Achieving equity in special education: history, status, and current challenges. *Exceptional Children, 74*(3), 264-288.

Tindal, G., & Nese, J. F. T. (2012, April). *Within year achievement growth using curriculum based measurement.* Paper presented at the National Council on Measurement in Education Vancouver, B. C., Canada.

Vallas, R. (2009). The disproportionality problem: The overrepresentation of black students in special education and recommendations for reform. *Journal of Sociology and Policy 17 VA*(181), 184.

Wang, C., Algozzine, B., Ma, W., & Porfeli, E. (2011). Oral Reading Rates of Second-Grade Students. *Journal of Educational Psychology, 103*(2), 442–454.

Zorigian, K., & Job, J. (Oct. 12, 2012). Gender in Special Education Retrieved Oct. 12, 2012, from http://www.learnnc.org/lp/pages/6817

Table 1. Comparison of Studies on Research Variables Referenced in Studying Growth of Reading Fluency

Researchers	**Grades**	**N**	**Measure**	**N of Meas.**	**Slope Calc.**	**Ave. Growth (per week)**
Fuchs, Deno, and Mirkin (1984)	Grades 3-5	64 Trt 77 Ctl	3rd grade passage reading test from Ginn 720	Unknown (from Nov. through May)	Pre-post difference @ 28 weeks	Treatment = 42-70 (1 wcpm) Control = 51-51 (0 wcpm)
Fuchs, et al. (1993)	1 – 6	16-25	'Generic passages were used' (p.31)	Unknown with 7+ measures for quadratic calculation	Ordinary Least Squares (OLS)	Grade 1 (n=19) @ 2.10 Grade 2 (n=25) @ 1.46 Grade 3 (n=14) @ 1.08 Grade 4 (n=16) @ .84 Grade 5 (n=20) @ .49 Grade 6 (n=23) @ .32
Simmons, Fuchs, Fuchs, Mathes, and Pate (1995)	Grades 2-5	21 (ET) 18 (ETPT) 29 (CTL)	Comprehensive Reading Assessment Battery (folktales)	16 weeks Explicit Teaching (ET), ET + Peer Tutoring (PT), and a Control Condition	Pre-post difference	
Fuchs, Fuchs, Mathes, and Simmons (1997)	Ave. 3rd grade	25 LD 21 Low 20 Ave	Comprehensive Reading Assessment Battery (folktales)	15 weeks of Peer Assisted Learning (PALS) or control	Pre-post difference	For PALS treatment 3.4 wcpm for LD 3.13 wcpm Low 4.00 wcpm Ave
Deno et al. (2001)	1 – 6	2,999	'Grade-appropriate' as determined by LEAs	Weekly and seasonally (fall, winter, spring)	Ordinary Least Squares (OLS)	Grade 1-2 @ 1.82 (GE) Grade 1-2 @ .71 (SE) Grade 3-4 @ 1.11 (GE) Grade 3-4 @ .58 (SE) Grade 5-6 @ .62 (GE) Grade 1-2 @ .60 (SE)

Schatschneider, Wagner, and Crawford (2008)	Grade 1	23,438 students attending Reading First schools	DIBELS oral reading fluency and SAT 10	The ORF was administered four times during the months of September, December, February, and April.	Compare predictive validity of estimates of (a) student growth in oral reading fluency, (b) student status or level of oral reading fluency, and (c) combined measures of growth and status, for prediction of concurrent and future reading skills	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. Period / M / SD Sep 13.68 18.91 Dec 20.33 22.40 Feb 34.27 27.75 Apr 49.55 32.26
Puranik, Petscher, Al Otaiba, Catts, & Lonigan (2008)	Grades 1-3	1,991 participants, of whom 1,388 with SI and 603 with LI in first grade; divided into resolved and persistent with 6 subgroups.	Oral Reading Fluency measure from the Dynamic Indicators of Early Literacy Skills (DIBELS; Kaminski & Good, 1996)	Four times per year (i.e., September, December, February, April)	Hierarchical piecewise growth curve model (PGCM) across first, second, and third grades	Average (WCPM/month) *Gr1* / *Gr2* / *Gr3* SI-PE .49 -2.13 5.11 LI-PE -.11 -1.90 4.60 SI-RE .89 -1.89 5.09 LI-RE 1.02 -3.66 3.99 SI-LD .42 -2.44 4.49 LI-LD .11 -1.58 3.72 Norm .49 -1.81 5.45 Language- impaired subgroups showed poorer performance com-pared to speech-impaired subgroups
Ardoin & Christ (2008)	K – 5	540	DORF 2nd grade passages	Three universal screenings	Gain from fall to spring	.99 to 1.04 words correct per week growth
Graney, Missall, Martinez, & Bergstrom (2009)	3 – 5	442 Yr1 456 Yr2	R-CBM AIMSweb passages from 2004-2006	Screening passages in fall, winter, and spring	Raw score difference / number of weeks in interval	.55 (F-W) vs .94 (W-S) in year 1 .92 (F-W) vs 1.12 (W-S) in year 2
Jenkins, Graff, & Miglioretti (2009)	4 – 8	41 LD	Standard reading passages from Vanderbilt	A total of 29 measures in 1, 2, 3, 4 weeks and pre-post (Sept.-Nov.)	Linear regression	1.09 word correct (true slope) 1.49 word correct (1 BL-1 wk) 1.94 word correct (1 BL-2 wk) 1.77 word correct (1 BL-3 wk) 1.83 word correct (1 BL-4 wk) 1.60 word correct (1 BL-pre-post)

Al Otaiba, Petscher, Pappamihiel, Williams, Dyrlund, & Connor (2009)	Grades 2- 3	5,004 Latino students (a) proficient in English, (b) not proficient and receiving English as a second language (ESL) services, and (c) proficient enough to have exited from ESL	Oral reading fluency from Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 1996) and PPVT-III	Four assessment periods: first 20–30 days of school (Sep); between the 65th and 75th days of school (Nov); between the 110th and 120th days of school (Feb); and between the 155th and 165th days of school (Apr).	Two level HLM with growth centered on the first testing time in third grade (September) and student characteristics (i.e., language group and special education subgroup) were entered at Level 2 designed to model both second- and third-grade growth trajectories .	Grade 2 Average WCPM/week *Latino* *ESL* *ESL-exit* Gen Ed 1.19 .76 1.16 LD 1.23 .75 1.08 SL 1.25 .93 1.08 Grade 3 Average WCPM/week *Latino* *ESL* *ESL-exit* Gen Ed 1.23 1.10 1.29 LD 1.31 1.10 1.28 SL 1.23 1.08 1.19
Christ, Silberglitt, Yeo, & Cormier, (2010)	2 – 6	4,824	AIMSweb CBM-R passages from 2001-2005	Fall, winter, and spring passages (0, 18, and 36 weeks)	Linear Mixed Model (LMM) for linear and piece wise growth	Growth of .88-1.71 vs .74-1.02 (GE) Growth of .69-1.17 vs .73-1.08 (SE)
Crowe, Connor, & Petscher (2009)	Grade 1 (9,993), grade 2 (9,869), and grade 3 (10,141)	Lower SES (eligible for free or reduced price lunch) and non lower SES students	Grade level passages from Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002).	Four measures in September, December, February, and April	Hierarchical Linear Modeling (HLM) to estimate mean growth trajectories for curriculum interacting with SES over the school year (7 months)	Linear growth in grades 1 (~15-50) and 2 (~50-80) reflecting modest acceleration and deceleration of growth in grade 3 (~70-90). Curriculum differences were reported between Reading Mastery, Scott Foresman, Harcourt, Success for All, Open Court, and Houghton Mifflin, which also interacted with students from different SES groups.

Logan & Petscher (2010)	Grades 1-3	Grade 1 (58,844), grade 2 (56,768), and grade 3 (57,873)	Grade level passages from Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002).	Four measures in September, December, February, and April in 2005-2006 academic school year	Three-level growth curves were fit and latent profile analysis used to cluster schools into groups based on their percentages of students meeting three risk categories (percentage of minority, ELL, and FRL students within each school) with six models were compared	Four different profiles or groups of schools existed in the data: Low-Risk (starts at 18 with 4.01 WCPM/month growth), Average-Risk (starts at 17 with 3.45 WCPM/month growth), Poverty-Risk (15 with 3.15 WCPM/month growth), and Language-Risk (starts at 14 with 3.08 WCPM/month growth)
Kim, Petscher, Schatschneider, & Foorman (2010)	Grades 1-3	13,154 enrolled in a Reading First school over 4 years	Phoneme segmentation, letter names, nonsense word fluency, and oral reading fluency from Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002); SAT-10 comprehension and PPVT-III	Fall, winter, spring in grades 1-3	Five multilevel growth models were used to estimate students' average level of performance at each grade (i.e., initial or end-of-year status), as well as the average rate of change for each of the predictors on their reading comprehension achievement; dominance analysis was used to compare 31 sets of predictors	*1st Gr M (SD)= 3 WCPM/mon* Fall – 22.5 (19.9) Winter1 – 35.1 (25.6) Winter2 – 45.5 (30.1) Spring – 55.5 (30.4) *2nd Gr M (SD)= 4.4 WCPM/mon* *Fall – 59.2 (31.1)* Winter1 – 72.8 (30.1) Winter2 – 85.8 (33.6) Spring – 96.9 (33.6) *3rd Gr M (SD)= 5.6 WCPM/mon* Fall – 86.1 (30.6) Winter – 101.1 (31.3) Spring – 121.3 (34.5) 1st gr ORF growth most predictive
Jenkins, J., & Terjeson, K. (2011)	2 – 6	31	Std. reading passages 1-2 grades below student grade level	Measures every 2, 4, and 8 weeks	Least squares slopes	1.67 (every 2 weeks) 1.48 (every 4 & 8 weeks) 1.29 (grade 3 @ 8 weeks) 1.63 (grade 5 @ 8 weeks)

Wang, Algozinne, & Ma (2011)	Grade 2	5,796 second-grade students in a large urban public school system in North Carolina.	Oral reading fluency from Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Good & Kaminski, 2002)	ORF was measured three times (fall, winter, and spring) during the second grade	Growth curve analysis was completed using Hierarchical Linear Model (HLM)	General Ed / Male / Female Fall 61.08 69.53 Winter 87.87 97.49 Spring 103.27 111.60 Special Ed / Male / Female Fall 47.18 51.87 Winter 68.61 76.28 Spring 84.75 89.34
Keller-Margulis, Clemens, Im, Kwok, & Booth (2012)	Grades 3-5	Grade 3 (1838), Grade 4 (2151), and grade 5 (2332) ELL, ELL-Monitor, and Non-ELL	Grade-level R-CBM probes from AIMSweb (2012)	Three measures in September, January (14 weeks later), and May (14 weeks later).	A piecewise model was fitted to the data that modeled growth	Found a seasonal effect in 3^{rd} & 4^{th} grade and with 3 groups. Table below is for successive grades (within cells) and pieces 1 and 2 (across rows). ELL / ELLM / Non ELL 19, 15, 15 / 20, 17, 14 / 20, 17, 15 18, 12, 13 / 15, 15, 15 / 15, 13, 13
Nese, Biancarosa, Anderson, Lai, Alonzo, & Tindal (2012)	3 – 5	2,465	easyCBM passages	Benchmark passages (fall, winter, spring)	Hierarchical Linear Model (HLM) for linear and discontinuous growth	Grade 3 – curvilinear (74 -106 -108) Grade 4 – curvilinear (99 – 123 -131) Grade 5 – linear (132 – 142 -156)

Table 2. *Grade 3 descriptive statistics for sample used in both levels*

Variable	*N*	Mean	*SD*	Min	Max
Sex	614	0.52	0.50	0	1
Disability	614	0.21	0.41	0	1
Ethnicity	614	0.42	0.49	0	1
ELL	614	0.13	0.33	0	1
Grade Level	614	0.18	0.38	0	1
PRFs	614	3.56	3.12	1	19

Table 3. *Grade 4 descriptive statistics for sample used in both levels*

Variable	*N*	Mean	*SD*	Min	Max
Sex	959	0.54	0.50	0	1
Disability	959	0.22	0.41	0	1
Ethnicity	959	0.52	0.50	0	1
ELL	959	0.14	0.34	0	1
Grade Level	959	0.14	0.35	0	1
PRFs	959	2.88	2.93	1	30

Table 4. Grade 5 descriptive statistics for sample used in both levels

Variable	*N*	Mean	*SD*	Min	Max
Sex	456	0.55	0.50	0	1
Disability	456	0.31	0.46	0	1
Ethnicity	456	0.44	0.50	0	1
ELL	456	0.09	0.28	0	1
Grade Level	456	0.27	0.44	0	1
PRFs	456	4.55	3.32	1	24

Table 5. Grade 3 – Unconditional model with intercept and slope

Fixed Effect	Coefficient	Std. Error	T-ratio	d.f.	P-value
For intercept	72.03	1.76	40.94	613	<0.001
For slope	0.67	0.04	16.28	613	<0.001

Table 6. *Grade 3 – Conditional model with student characteristics and measurement conditions*

Fixed effect	Coefficient	*SE*	t-ratio	*df*	*p*-value
Intercept	99.53	3.11	32.03	607	<0.001
Sex	-5.96	3.06	-1.95	607	0.052
Disability	-22.61	4.50	-5.03	607	<0.001
Ethnicity	0.78	3.30	0.24	607	0.812
ELL	-9.07	5.43	-1.67	607	0.095
Grade Level	-13.83	4.33	-3.19	607	0.002
PRFs	-4.01	0.38	-10.47	607	<0.001
Slope	0.21	0.10	2.16	607	0.031
Sex	0.15	0.08	1.95	607	0.052
Disability	0.01	0.09	-0.11	607	0.916
Ethnicity	0.05	0.08	0.69	607	0.489
ELL	-0.03	0.14	-0.18	607	0.861
Grade Level	-0.14	0.10	-1.39	607	0.166
PRFs	0.07	0.01	7.28	607	<0.001

Table 7. Grade 4 – Unconditional model with intercept and slope

Fixed Effect	Coefficient	Std. Error	T-ratio	d.f.	P-value
For intercept	97.10	1.36	71.38	958	<0.001
For slope	0.65	0.04	17.52	958	<0.001

Table 8. *Grade 4 – Conditional model with student characteristics and measurement conditions*

Fixed effect	Coefficient	*SE*	*t*-ratio	*df*	*p*-value
Intercept	115.61	2.70	42.83	952	<0.001
Sex	-0.93	2.47	-0.38	952	0.707
Disability	-25.32	3.71	-6.82	952	<0.000
Ethnicity	-1.56	2.58	-0.61	952	0.545
ELL	-23.83	3.90	-6.11	952	<0.001
Grade Level	-11.04	4.13	-2.67	952	0.008
Performances	-2.52	0.38	-6.60	952	<0.001
Slope	0.72	0.08	8.87	952	<0.001
Sex	-0.10	0.07	-1.34	952	0.181
Disability	-0.07	0.10	-0.72	952	0.472
Ethnicity	0.01	0.08	0.15	952	0.879
ELL	0.26	0.13	1.98	952	0.047
Grade Level	0.02	0.10	0.24	952	0.809
PRFs	0.00	0.01	0.56	952	0.579

Table 9. Grade 5 – Unconditional model with intercept and slope

Fixed Effect	Coefficient	Std. Error	T-ratio	d.f.	P-value
For intercept	103.80	1.67	62.20	455	<0.001
For slope	0.79	0.04	22.54	455	<0.001

Table 10. *Grade 5 – Conditional model with student characteristics and measurement conditions*

Fixed effect	Coefficient	*SE*	*t*-ratio	*df*	*p*-value
Intercept	121.90	3.18	38.33	449	<0.001
SEX	-3.68	2.81	-1.31	449	0.191
Disability	-28.32	3.22	-8.80	449	<0.001
Ethnicity	4.94	2.85	1.74	449	0.083
ELL	-7.08	4.90	-1.45	449	0.149
Grade Level	-17.70	3.53	-5.02	449	<0.000
PRFs	-0.90	0.42	-2.12	449	0.034
Slope	0.96	0.08	12.31	449	<0.001
Sex	-0.05	0.07	-0.65	449	0.517
Disability	0.05	0.08	0.67	449	0.504
Ethnicity	-0.05	0.08	-0.69	449	0.49
ELL	0.06	0.11	0.57	449	0.57
Grade Level	-0.13	0.08	-1.62	449	0.106
PRFs	-0.02	0.01	-2.29	449	0.022