School Psychology Review,
2010, Volume 39, No. 3, pp. 447–462

Curriculum-Based Measurement of Oral Reading: An Evaluation of Growth Rates and Seasonal Effects Among Students Served in General and Special Education

Theodore J. Christ
University of Minnesota

Benjamin Silberglitt
Technology and Information Educational Services

Seungsoo Yeo
Korean Educational Development Institute

Damien Cormier
University of Minnesota

Abstract. Curriculum-based measurement of oral reading (CBM-R) is often used to benchmark growth in the fall, winter, and spring. CBM-R is also used to set goals and monitor student progress between benchmarking occasions. The results of previous research establish an expectation that weekly growth on CBM-R tasks is consistently linear throughout the academic year. The patterns of CBM-R growth were examined for a large sample of students (N = 3808) from both general education and special education populations in second to sixth grades. Results support four general conclusions: (a) annual growth is more substantial within the general education population; (b) growth is more substantial in earlier elementary grades; (c) more growth occurs in the fall than the spring season (i.e., seasonal effect), especially within the early primary general education population; and (d) the seasonal effect is less pronounced within the special education population. Estimates of growth within and across seasons are presented and implications are discussed.

Curriculum-based measurement (CBM) is used to index annual student growth across the primary grades. Procedures and measurement metrics are developed for mathematics, spelling, written expression, and reading. CBM oral reading (CBM-R) rate is the most researched and well established of those available procedures (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). CBM-R is used as an index of reading rate and fluency, which are identified as critical skills to target for instruction and intervention within the early stages of

Correspondence regarding this article should be addressed to Theodore J. Christ, 56 East River Road, Minneapolis, MN 55455; E-mail: tchrist@umn.edu

reading development (National Reading Panel, 2000; National Research Council, 1998). CBM-R outcome also functions as a robust indicator of overall reading development throughout the primary grades (Wayman et al., 2007).

CBM-R progress monitoring data are collected and used to evaluate instructional effects, determine when to modify the instruction, and evaluate response to intervention (Deno, 1985, 1986). Teachers who use progress monitoring data are likely to use more specific and measurable goals, rely more substantially on data to guide instruction, and modify instruction more frequently (Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, & Stecker, 1989). Progress monitoring applications of CBM-R have the potential to confer the most substantial benefit to students. Progress monitoring can be conceptualized as either interim benchmark/screening assessments when data are collected three to four times per year, or time series continuous assessments when data are collected daily or weekly.

Both interim and time series CBM-R data are plotted graphically and evaluated against goals and goal lines. Fuchs and Shinn (1989) recommended that the rate of expected growth should derive from local normative performance so that individual student achievement can be compared against the local normative performance of grade-level peers. A goal line is established by connecting the CBM-R data point for the observed level of performance from Week 1 (initial performance) to the expected/typical level of performance at Week 36 (ending performance). Once graphed, the goal line provides a graphic trajectory of expected growth for the academic year. Ongoing time series data are then plotted on the graph and evaluated against the goal line. For example, if a student's initial level of CBM-R performance in September was 20 words read correctly per minute (WRCM) and in May that student was expected to read 60 WRCM, then the expected annual growth is 40 WRCM across 8 months, or 32 weeks. That translates to ~1.25 WRCM per week of growth (1.25 WRCM = 40 WRCM/32 weeks). The goal line and trajectory of expected growth is based on the assumption of a monotonic and linear trend.

Monotonic Linear Growth

Standards of expected performance within and across the year provide the foundation to establish instructional goals and goal lines. There are two influential studies that established standards of weekly CBM-R growth (Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993), which have application in both research and practice. Fuchs et al. (1993) analyzed a CBM-R data set comprised of 103 students in general education and 14 students in special education. Each student was assessed weekly for an entire academic year using CBM-R procedures. Although some students evidenced a negatively accelerating annual growth, the researchers concluded that annual growth is, on average, monotonic and linear. They also observed a negatively accelerating trend across grades. That is, the rate of CBM-R growth was positive in all grades, but the rate of growth declined in each successive grade. The resulting analysis yielded grade-specific estimates of standard and ambitious rates of growth—and *assumed* a pattern of monotonic linear growth. Respectively, the growth, determined by the number of WRCM, were 1.5 and 2.0 in first and second grade, 1.0 and 1.5 in third, 0.85 and 1.1 in fourth, 0.5 and 0.8 in fifth, and 0.3 and 0.63 in sixth.

Deno et al. (2001) analyzed a large data set comprised of 2675 students in general education and 324 in special education classes. The student sample was from school districts across the country. The researchers relied on the *assumption* of monotonic linear annual growth. Ordinary least squares regression was used to calculate weekly growth. The study yielded useful estimates of observed and expected growth across grades and populations. The authors estimated that beginning readers in the general education population can be expected to improve 2 WRCM/week until they achieve 30 WRCM; thereafter, students in the general education population can be expect to improve at least 1 WRCM/week. Researchers observed that students in the special education population demonstrated substantially less growth, which approximated 0.5– 0.8 WRCM/week, unless they are provided with robust evidence-based instruction, which might improve the observed rate of growth to 1.39 WRCM/week.

Prior studies provide useful estimates of typical, ambitious, and expected rates of CBM-R growth; however, those estimates depend on the assumption that growth is monotonic and linear across the academic year. That assumption should be examined more closely to determine its veracity. The prior studies also rely on analyses that neglect the inherent nesting of students within classrooms and schools, which establish potential bias in the linear and quadratic functions derived.

Improved Modeling of Annual and Seasonal Growth

Research is necessary to evaluate alternate models of annual growth and test the predominant assumption that annual growth is monotonic and linear. This study will evaluate both linear (constant) and piece wise (nonconstant) models of growth from fall to winter to spring; moreover, this study will employ the superior approach of linear mixed model (LMM) rather than that of ordinary least squares used in prior research (Deno et al., 2001; Fuchs et al., 1993).

Ordinary least squares only estimates the fixed effects, which assumes that all cases have identical parameter estimates for intercept and growth. This is a limitation when modeling CBM-R growth because it is likely to be an erroneous assumption in most cases. The results of published research provide evidence for substantial magnitudes of variance in both intercept and growth among students in a sample (Shin, Deno, & Espin, 2000; Stage, 2001). It is likely that the intercept and growth rates should be treated as random effects in many instances. LMM can be used to estimate both the fixed effects and the random effects, which are the intercepts and growth rates for individual student cases in the sample. LMM takes into account the variances within and across individuals in the sample. Finally, ordinary least squares requires that missing data are removed casewise whereas LMM can handle missing data, which is likely to yield less biased estimates of intercept and slopes while enhancing power (Duncan, Duncan, & Strycker, 2006). This approach facilitates analysis of both seasonal effects, as pursued by two previous studies (Ardoin & Christ, 2008; Graney, Missall, Martinez, & Bergstrom, 2009), and compares those effects to that of linear growth as assumed in other studies (Deno et al., 2001; Fuchs et al., 1993).

The calendar year was divided into three seasons for the purpose of this study. The seasons were defined by those periods between standard tri-annual assessment occasions: fall, winter and spring. The duration between fall and winter assessments defines the fall season, that between winter and spring defines the spring season, and that between spring and fall defines the summer season. If growth is consistent and linear throughout the calendar year, then there is no *seasonal effect*. Conversely, if growth is not consistent and linear throughout the calendar year, then there is a seasonal effect whereby there is more growth observed for some seasons than for others.

The summer season is the most distinct for the majority of students because they are not exposed to formal instruction. The rate of achievement is likely to decline in the summer months (Kim, 2004) unless formal instruction is provided through summer school programs (Stage, 2001). A seasonal effect in the summer is expected because the conditions for learning are substantially different from that which occurs in the spring and fall seasons. The effect, however, does not seem specific to those differences between summer and the academic year. Ardoin and Christ (2008) observed statistically significant differences in CBM-R growth between the fall and spring seasons. That is, among a sample of second-grade students there was more growth in the fall than in the spring season. That result replicated across three combinations of CBM-R passages sampled from the Dynamic Indicators of Basic Early Literacy Skills. A subsequent study aimed to replicate and extend those findings with a sample of third- through fifth-grade students across a variety of CBM curriculum domains (Graney et al., 2009). However, Graney et al. (2009) concluded that there was more CBM-R growth in the spring season than in the fall season, failing to replicate the work of Ardoin and Christ. Graney et al. speculated that the failure to replicate findings across studies might relate to the specification of data collection schedules, sample characteristics, issues of instrumentation, or the distribution of instructional intensity across the academic year. Although it is difficult to make a single attribution, it should be noted that the student sample used by Graney et al. derived from a school–university partnership program that included features to promote data use and evidence-based interventions. Indeed, that program might influence the pattern of student growth; however, additional research and replications are necessary to help determine whether specific patterns for seasonal effects should be expected.

There are limitations to the analytic procedures used in the two prior studies. Both generated estimates of growth from difference scores across seasons. That is, growth was estimated as the difference between performance at Time 1 and Time 2, which was then divided by the number of intervening weeks to yield an estimate of weekly growth. Although this method is intuitive, and it was useful for the preliminary work to examine seasonal effects, the method of difference scores to estimate change has long been criticized as an unreliable method of analysis (Cronbach & Furby, 1970) and may have contributed to the inconsistent findings. The analytic procedures used in this study improve on those of prior studies that examined seasonal effects and modeled annual growth. The analytic procedures used in this study will compare the hypothesized seasonal effect directly with that of linear growth.

Purpose

The purposes of this study were to (a) estimate the rate of growth within and across seasons for general education and special education populations; (b) evaluate whether the rate of CBM-R growth is likely to be consistent (i.e., linear model) or inconsistent (i.e., piecewise model) throughout the academic year; and (c) test the hypothesis that the seasonal effect is typically characterized by more robust growth in the fall than for the spring season. That hypothesis is counter to the assumptions inherent to the linear models as applied in many previous studies and it is counter to the findings of Graney et al. (2009).

Method

Participants and Setting

A total of 4824 students (52% male, 48% female) in the second through sixth grades were included in the sample. The ethnicity breakdown of the sample was 94% White, not Hispanic; 2% Native American; 2% Asian or Pacific Islander; 1% Hispanic; and 1% Black, not Hispanic. Participants were enrolled in seven elementary schools, located within five school districts in the rural and suburban Midwest. The sample consisted of 31% students receiving free or reduced-price lunch, and 8% students receiving special education services.

As part of an ongoing benchmark assessment plan and response to intervention implementation (see Bollman, Silberglitt, & Gibbons, 2007, for additional information), all students in participating schools were assessed using CBM-R every fall, winter, and spring of every year across the grade levels studied. All students for whom relevant data were collected participated in the study. Data were gathered between the years of 2001 and 2005 for each student in the sample. Data from across multiple years were available for many of the students in the sample; thus many students participated in the study at multiple grade levels.

Measures

The CBM-R measures were developed or selected to be curriculum-independent, grade-level appropriate, and of equivalent difficulty across probes within each grade level. Each probe contained approximately 250 words typed on an 8 by 10 piece of paper in a grade-appropriate font (range = 12–14 point with larger print for the second and third grades). Probes were selected from a standardized set of commercially available published passages (AIMSweb; Howe & Shinn, 2002). The technical manual for AIMSweb passages (Howe & Shinn, 2002) and peer-reviewed research (Wayman et al., 2007) provide evidence that test–retest reliability and alternate-form reliabilities approximate or exceed .90. Criterion-related validity coefficients typically approximate the .70–.90 range. The same set of CBM-R probes were used across all districts for all 4 years of the study.

Procedure

CBM-R data were collected in 4-month intervals in the fall, winter, and spring of each academic year. Standardized administration and scoring procedures were implemented for each administration, which was conducted by trained school-based personnel and high school honors students. These individuals were trained to administer three successive CBM-R probes and score WRCM using standardized procedures that have appeared in the published literature (Howe & Shinn, 2002; Shinn, 1988, 1989). Administration occurred outside of the classroom to minimize distractions, in a separate room or quiet hallway while the administrator and student were seated at desks or in chairs.

Training included a brief instructional session that lasted approximately 1 hr. Each training session was followed by an assessment of competency that included practice and evaluation of scoring procedures. No formal criterion for scoring was established until the final years of the study. In 2003–2004 and 2004–2005, each scorer was required to come within 2 WCPM of the correct score on three consecutive videotaped assessments. Although data were not maintained, few scorers required more than one opportunity to meet criterion. Finally, inter-rater reliability was not evaluated; however, CBM-R procedures have established high levels of inter-rater reliability that appear consistently within the published literature. A survey of 10 arbitrarily selected CBM-R studies that were published between 1995 and 2005 reported inter-rater reliabilities that were consistently above .95.

Analytical Plan

A LMM analysis was performed to investigate evidence of seasonal effects for CBM-R. The present LMM analysis considered two hypothesized growth models: linear and piecewise (Duncan et al., 2006; Flora, 2008). The linear model assumes a static growth rate across time points—meaning that regardless of change over time, the estimated growth rate from the linear model is identical to the sample's growth rate. In contrast, the piecewise model assumes that the developmental trajectory is not consistent across time, making it possible to provide two different linear trends. Thus, the piecewise model takes into account a possible nonlinear trajectory for CBM-R. Based on the three CBM-R data points used in this study, the piecewise model can be parameterized into two levels of analyses. First, Level 1 can be expressed as

$$y_{it} = \beta_{0i} + \beta_{1i}\,\lambda_{1t} + \beta_{2i}\lambda_{2t} + \varepsilon_{it} \quad (1)$$

where y_{it} is the outcome of CBM-R for subject i at time t, β_{0i} is the value of CBM-R for subject i in fall, β_{1i} is the first linear trend of subject i from fall to winter, λ_{1t} is the value of the linear trend at time t (Week 0 and Week 18), β_{2i} is the second linear trend of subject i from winter to spring, λ_{2t} is the value of the linear trend at time t (Week 18 and Week 36), and ε_{it} is unique measurement error for subject i at time t.

Level 2 can be described as

$$\beta_{0i} = \beta_0\zeta_{0i} \quad (2)$$

$$\beta_{1i} = \beta_1\zeta_{1i} \quad (3)$$

$$\beta_{2i} = \beta_2\zeta_{2i} \quad (4)$$

where β_0 represents the average initial status for all subjects, β_1 refers to the average of the first linear growth rate for all subjects, and β_2

refers to the average of the second slope trend for all subjects. ζ_{0i}, ζ_{1i}, and ζ_{2i} refer to the random deviation of the initial status, the first growth rate, and the second growth rate, respectively. As a result, the piecewise model presented earlier assumes that the second slope is statistically different from the first slope, indicating that there is a seasonal effect for CBM-R throughout the academic year.

The first step in this analysis was to compare the two models (linear and piecewise models) using fit indices, which included the likelihood ratio test (LRT), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The three fit indices are considered to be "goodness-of-fit" indices, where the smallest fit statistics are viewed as the best model. If results from these goodness-of-fit indices reveal that the linear model is better than the piecewise model, then the slope is the same from fall to spring and there is no evidence of a significant seasonal effect for CBM-R across the academic year. Conversely, if goodness-of-fit indices are better for the piecewise model than the linear model, then we conclude that the slopes are different from fall to winter and from winter to spring and there is evidence of a seasonal effect across the academic year. If the piecewise model is the better fit, then an additional model is needed to estimate the extent to which the first slope differs from the second slope.

The additional model, which is called *added growth model,* would be identical to the piecewise model described earlier, except for the supplementary time factor loading matrix (Flora, 2008). For example, in the piecewise model, the time factor loading can be described as

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 18 & 18 \end{bmatrix} \begin{matrix} \textit{Week} & 0 \\ \textit{Week} & 18 \\ \textit{Week} & 36 \end{matrix} \qquad (5)$$

where the first column is set at 1, meaning that the outcome of the fall CBM-R score (intercept) does not change across time. The second column refers to the first linear time metric indicating change in CBM-R score from fall to winter (Flora, 2008). Similarly, the third column represents change in CBM-R score from winter to fall.

For the added growth model, the time factor loading can be expressed as:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 18 & 0 \\ 1 & 36 & 18 \end{bmatrix} \begin{matrix} \textit{Week} & 0 \\ \textit{Week} & 18 \\ \textit{Week} & 36 \end{matrix} \qquad (6)$$

The matrix in Equation 6 is identical to the matrix in Equation 5, with the exception of the second column with 0, 18, and 36. The different time factor loadings for the additional model is applied to examine the extent to which linear change in CBM-R from fall to winter increases or decreases compared to that from winter to spring. In other words, the added growth model analysis makes it possible to contrast the first linear slope and the second linear slope. The results from the additional analysis can also be viewed as an effect size estimate between the first and second linear slope (Flora, 2008).

Data Preparation

Overall, the rates of missing data for fall, winter, and spring seasons were less than 4% for all grade levels. All LMM analyses were conducted using the SAS MIXED program with full maximum likelihood estimation that automatically conducts all adjustments for missing data. Monthly CBM data used in this study were nested within districts. Therefore, a three-level LMM was used to satisfy an assumption that data collected in this study might not be independent among school districts. An alpha level for all statistical tests in this study was set at .001.

Model assumptions were evaluated, which included a test of normality using Kolmogorov-Smirnove statistics, skew, and kurtosis, and homoscedasticity of residuals by plotting raw residuals at each level (Levels 1 and 2). No violations were detected. The Level 3 LMM analysis was used to test whether CBM data used in the study were nested within school districts. As a result, variations in the slopes and intercepts between school districts were not significantly different

Table 1
Descriptive Statistics for the Level and Change of CBM-R

	Seasonal Benchmark Levels			Weekly Change	
	Fall	Winter	Spring	Fall Season (Fall–Winter)	Spring Season (Winter–Spring)
Population	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
General education					
Second	57 (32)	88 (34)	107 (35)	1.71 (.73)	1.02 (.64)
Third	82 (37)	107 (38)	125 (40)	1.38 (.79)	0.97 (.75)
Fourth	106 (38)	128 (39)	141 (41)	1.19 (.72)	0.78 (.70)
Fifth	125 (40)	145 (41)	160 (42)	1.12 (.76)	0.83 (.75)
Sixth	142 (41)	158 (42)	171 (43)	0.89 (.71)	0.74 (.75)
Total	99 (47)	122 (45)	137 (46)	1.26 (.74)	0.87 (.72)
Special education					
Second	33 (28)	55 (35)	74 (38)	1.17 (.81)	1.08 (.63)
Third	52 (33)	72 (39)	88 (41)	1.09 (.81)	0.89 (.78)
Fourth	77 (40)	96 (43)	108 (45)	0.99 (.72)	0.67 (.65)
Fifth	89 (43)	106 (45)	118 (47)	0.90 (.76)	0.67 (.79)
Sixth	106 (47)	119 (50)	132 (51)	0.70 (.72)	0.73 (.69)
Total	72 (45)	89 (48)	104 (49)	0.97 (.77)	0.81 (.71)

Note. CBM-R = curriculum-based measurement of oral reading. Skew and kurtosis were within reasonable limits ($z \leq 2.00$) for all conditions. *N* values for second, third, fourth, fifth, and sixth grades were, respectively, 1790, 1885, 1913, 1942, and 1144 for general education; 111, 134, 170, 192, and 116 for special education.

from zero for all grade levels, meaning that both the slopes and intercepts at the district levels were identical. In addition, such a result supports evidence that the Level 2 LMM analysis introduced in the method section is appropriate for this study.

Results

Summaries of the means and standard deviations for CBM-R at each time point are presented in Table 1. As shown in Table 1, the average growth as measured with CBM-R was positive for all conditions across seasons (fall–winter, winter–spring), grade levels (second to sixth) and populations (general education, special education).

Comparison Between Linear Model and Piecewise Model

Results for the comparisons of linear and piecewise models are presented in Table 2. For students in general education, the piecewise model provided the best fit for tri-annual assessment data, as evidenced by the magnitudes of LRT, AIC, and BIC observed for the linear and piecewise models respectively. The χ^2 difference test between the two models also revealed that the piecewise model fit significantly better than the linear model.

For students in special education, on the basis of the LRT criterion, the piecewise model for all grade levels was viewed as the best-fitting model. Significant differences in LRTs between the two models, however, were found in the second through fourth grades. There were inconsistencies with respect to AIC and BIC criteria, which might indicate the linear model fit better in some cases within the sample of special education students. Although the results of AIC and BIC are noted, the LRT was the primary test/criterion used here.

Table 2
Deviation Between Linear and Piecewise Models for General Education Students and Special Education Students

Grade	Linear Model			Piecewise Model			Differences (Linear-Piecewise Model)		
	LRT	AIC	BIC	LRT	AIC	BIC	LRT	AIC	BIC
General education									
Second	46231.7	46239.7	46261.7	45542.6	45556.6	45595.0	689.1*	683.1	666.7
Third	49566.9	49574.9	49597.0	49357.7	49371.7	49410.5	209.2*	203.2	186.5
Fourth	49768.1	49776.1	49798.3	49520.1	49534.1	49573.0	248.0*	242.0	225.3
Fifth	50974.9	50982.9	51005.2	50866.6	50880.6	50919.6	108.3*	102.3	85.6
Sixth	29892.9	29900.9	29921.1	29873.6	29887.6	29922.9	19.3*	13.3	−1.8
Special education									
Second	2816.3	2824.3	2835.1	2802.6	2816.6	2835.6	13.7*	7.7	−0.5
Third	3501.8	3507.8	3516.5	3487.9	3495.9	3507.5	13.9*	11.9	9.0
Fourth	4394.9	4402.9	4415.4	4378.9	4392.9	4414.9	16.0*	10.0	0.5
Fifth	5090.6	5098.6	5111.6	5083.3	5097.3	5120.1	7.3	1.3	−8.5
Sixth	3048.2	3056.2	3067.2	3046.3	3060.3	3079.5	1.9	−4.1	−12.3

Note. LRT = likelihood ratio test; AIC = Akaike information criterion; BIC = Bayesian information criterion.
$^*p < .001$.

Comparison Between the First Slope and Second Slope for General and Special Education

Table 3 presents estimates of both the first (fall–winter) and second (winter–spring) linear slope. In the cases of both students in general education and special education, all of the seasonal linear slopes were positive with magnitudes significantly different than zero ($p < .001$). That observation was consistent with the conclusion that reading achievement accelerated continuously throughout the academic year for all student samples. Although the mean rate of growth across seasons was consistently greater for students in general education as compared with students in special education, it was the first slopes that were most distinct. Further, the differences in the second slopes varied by grade level.

Additional analysis was conducted to examine whether the first linear slopes (fall) were significantly different from the second slopes (spring). Table 3 presents the differences between the two linear slopes along with the results of significance tests. Figure 1 depicts the slopes and model differences.

In the sample of general education students, statistically significant differences ($p <$ .001) were observed between the first and second slopes in all grade-level conditions (Table 3). That is, the rate of growth for all grade levels dropped in the second part of the academic year. It is relevant to note that the magnitude of difference between the two slopes (seasonal effects) and standardized effect sizes consistently decreased as grade level increased. This might indicate that the seasonal effect in reading is larger in the lower primary grades than the upper primary grades. The power estimates for detecting seasonal effects for students in general education were very large for all grade levels (range = 0.96–1.00). This indicates a good chance of detecting a false null hypothesis (linear model) when the alternative model (piecewise model) is correct (Duncan et al., 2006).

In the sample of special education students, the average of the first slope (fall) was

Table 3
Comparison Between the First Slope and the Second Slope for General Education Students and Special Education Students

Grade	Intercept $\hat{\beta}_0$ (*SE*)	Fall/First Slope (Fall to Winter) $\hat{\beta}_1$ (*SE*)	Spring/Second Slope (Winter to Spring) $\hat{\beta}_2$ (*SE*)	Difference Between Fall Slope and Spring Slopes $\hat{\beta}_2^a$ (*SE*)	Standardized Effect Sizes[b]	Power
General education						
Second	57.42* (.76)	1.71* (.02)	1.02* (.01)	0.69* (.02)	34.50	1.00
Third	82.31* (.85)	1.38* (.02)	0.97* (.02)	0.41* (.03)	13.60	1.00
Fourth	105.78* (.93)	1.19* (.02)	0.78* (.02)	0.41* (.12)	3.41	1.00
Fifth	124.43* (.89)	1.12* (.02)	0.83* (.02)	0.29* (.12)	2.41	1.00
Sixth	142.49* (1.21)	0.88* (.02)	0.74* (.02)	0.14* (.16)	0.87	0.96
Special education						
Second	34.02* (2.67)	1.17* (.08)	1.08* (.06)	0.09 (.10)	0.90	0.86
Third	50.85* (2.81)	1.09* (.07)	0.89* (.06)	0.20 (.10)	2.00	0.87
Fourth	75.50* (2.99)	0.99* (.05)	0.67* (.05)	0.32* (.08)	4.00	0.92
Fifth	86.28* (3.06)	0.90* (.05)	0.67* (.06)	0.23 (.09)	2.50	0.54
Sixth	106.98* (4.40)	0.69* (.07)	0.73* (.06)	−0.04 (.10)	0.40	0.11

Note. *SE* = standard error.
[a]The estimated coefficients of the second slope from an added growth model.
[b]Standardized effect sizes = $\hat{\beta}_2$/SE, which is a strategy developed by Satorra and Saris (1985), was used to calculate power.
*$p < .001$.

greater than the second slope (spring) in all grades except for sixth; however, those differences were not statistically significant except in the fourth-grade condition (Table 3). Moreover, the pattern of reduced seasonal effects for progressively higher grades did not generalize to the sample of special education students. Instead, the magnitude was greatest in fourth grade (difference in slopes = 0.32, $p < .001$), somewhat smaller in third (0.20, $p > .001$) and fifth (0.23, $p > .001$), and hardly perceptible in second (0.09, $p > .001$) and sixth grades (−0.04, $p < .001$). The range of standardized effect sizes and power was variable across grades. Based on the results, it appears that, in general, the observed slopes in the sample of special education students were relatively consistent across seasons—in all conditions but fourth grade.

Discussion

Many of the published estimates of CBM-R annual growth rates rely on the assumption that the pattern is consistent and linear throughout the academic year (Deno et al., 2001; Fuchs et al., 1993). Although there is evidence for seasonal effects within some studies (Ardoin & Christ, 2009; Fuchs et al., 1993; Graney et al., 2009), the phenomenon of seasonal effects are generally ignored within the professional literature. This study yielded (a) estimates for the magnitude CBM-R annual growth along with those within the fall and spring seasons, and estimates were derived for both general education and special education populations; (b) both linear and piecewise growth models; and (c) hypothesis tests to examine whether fall growth is typically more substantial than spring growth.

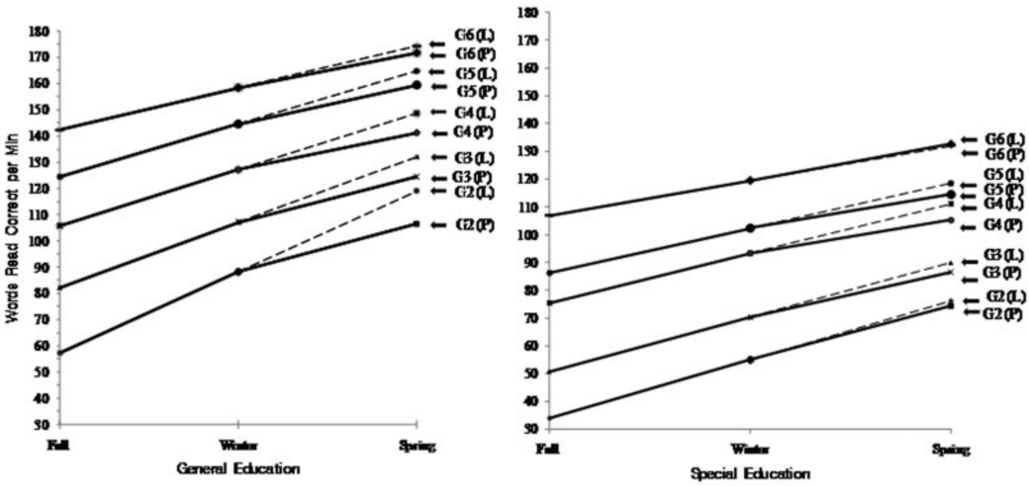

Figure 1. Differences for linear (L, solid line) and piecewise (P, dashed line) models of growth (G) for students in second through sixth grade within either the general (left) or special education population (right)

CBM-R growth was modeled and analyzed for five primary grades (second through sixth) and two student populations within and across seasons.

The results of this study support four general conclusions. Descriptive data make clear that, first, annual growth is greater among students in the general education population than for those in the special education population, especially in the fall season; and, second, more growth occurs in the early grades with less in the upper grades. Modeling and inferential analysis provide support for the third and fourth conclusions that follow. That is (third conclusion), there is a significant *seasonal effect* such that more growth occurs in the fall season than in the spring season for general education, but the magnitude of that effect declines with each progressive grade level. Finally, the seasonal effect is less pronounced among students in the special education population. In general, linear growth should not be assumed, seasonal effects should be expected—at least within the general education population—and piecewise models tend to fit better than linear models of growth.

Grade and Population Effects

Larger magnitudes of growth were observed for students in general education as compared to special education. This outcome was observed across five districts where CBM-R was used to screen, monitor, and evaluate instructional effects as part of a problem solving model. The systematic and ongoing use of CBM-R for these purposes did not establish equivalent growth rates between the student populations. Despite inconsistent growth rates across student groups, the rate of growth among students in the special education population was greater than observed in other studies (Deno et al., 2001; Fuchs et al., 1993). For example, the most recent estimates for average growth among students in the special education population were ~0.60 WRCM/week (Deno et al., 2001). That rate of growth was relatively consistent across grades. In contrast, the average growth estimates among students in the sample of students in special education used in this study were within the range of 0.67–1.17 WRCM/week (Table 3). Although cross-study comparisons cannot be used to establish that differences in systems-level practices (e.g., problem solving) and instructional practices were causally related to the observed differences in growth rates, future research should investigate the phenomena more closely.

The results of this study converge with previous research (Deno et al., 2001; Fuchs et al., 1993) to support the conclusion that there is more CBM-R growth within the early primary grades. This phenomenon was consistent for both fall and spring seasons in both the general education and special education populations (Table 1). There was a steady decline for both general education and special education in weekly growth as grade level was increased. There was a corresponding decline in the magnitude of seasonal effects as grade levels increased. Results also support the conclusion that seasonal effects are more robust for students served within general education and less robust for students served within special education.

Seasonal Effect

The primary purpose of this research was to determine whether the rate of CBM-R growth is consistent across the year. Notwithstanding limitations, a seasonal effect was sufficiently robust and consistent across grades and student populations to warrant serious consideration. Estimates of weekly CBM-R growth within season are reported in Table 1. They are also depicted in Figure 1. Both visual analysis and statistical analysis of growth models support the conclusion that there was a *difference in the average rate of growth across seasons*, and that difference occurred in all grades, but was generally more robust in the lower grades of general education (Figure 1). With few exceptions, more growth occurred in the fall than in the spring, which is depicted as the differences in the slopes within Table 3.

The *magnitudes of seasonal effects were inconsistent* across grades and student populations. In general education, the difference between fall and spring slopes ranged from 0.14 to 0.69 per week, and all the differences were statistically significant with substantial effect sizes (Table 3). That effect is illustrated in Figure 1, which reveals that the interactions between seasonal effect and student population were most distinct in second grade. That is, the seasonal effect was most robust for the sample of students in second-grade general education classes, with a difference in growth from fall to winter of 0.69, and minor in comparison for students in second-grade special education classes, with a difference of 0.09.

The results of this study support the conclusion that growth differs between fall and spring seasons, which is inconsistent with a number of currently held views in the field. The consistency of the results across grades for general education begs the question of why this effect may be occur and what are the potential implications. There are a number of alternate hypotheses that offer some insight on the phenomenon and directions for future research.

Explanatory Hypotheses

This study was designed to examine *whether* seasonal effects are likely to occur in cases of CBM-R tri-annual assessment and challenge the assumption of linear growth. Indeed, the results do provide support for seasonal effects; however, this study was not designed to examine the cause of seasonal effects. That is left to future research to isolate the cause(s) and potential remedies to accelerate reading achievement within seasons with deficit growth. The result of both this study and Ardoin and Christ (2008) converge to provide evidence that the seasonal effect is likely to coincide with deficit levels of growth in the latter part of the academic year, or the spring season, especially for the students in the early primary general education population. It seems that changes in instructional conditions, or student motivation, might explain deficit rates of growth within the spring season unless there is an active intervention to prevent those changes. It seems that data-based decision making could function to reduce, or reverse, the pattern of seasonal effects that were replicated across this study and that of Ardoin and Christ.

Teacher expectations. There are a number of ways in which teacher expectations can have an influence on student behavior, which might be conveyed through discrepant use of praise and punishment statements, content-relevant feedback, individualized attention, differentiated instruction, and variability in the opportunities to respond and engage in instructionally relevant practice (Good, 1981). Prior research yielded conclusions that there are indeed sometimes differences in teacher treatment of high- and low-achieving students over the course of the year, which affected the amount and type of instruction and feedback that students received at different points in the academic year (Good et al., 1980). Researchers concluded that teachers are more invested in influencing and shaping student behavior at the beginning of the school year than later in the year. Furthermore, the differentiation in teacher behavior may have sustained effects and become more influential as the school year progresses, resulting in lower achievement and growth later in the school year in low-achieving students. Other research also demonstrated the potential influence of teacher expectations, including expectation biases across ethnic groups (McKown & Weinstein, 2007) and ability groups (Rosenthal & Jacobson, 1966).

Classroom management. There is evidence that classroom management and organizational strategies are implemented with greater fidelity and effort in the early part of the academic year (Cameron, Connor, & Morrison, 2004). The same research provides evidence that that those educators who approach classroom management and organization with the greatest fervor early in the year are those same educators whose practices drop off most substantially later in the academic year. Improved levels of engagement and instructional gain are clearly associated with good classroom management and organization, and are among better predictors of student achievement (Brophy, 1987). Taken together, research provides a foundation to postulate that differential levels of classroom management and organization across seasons might influence student achievement. This effect is hypothesized to be the greatest at lower grade levels, as younger students require greater direction in organization practices (Cameron et al., 2004). The pattern of seasonal effects observed in this study coincides with a classroom management and organizational hypothesis.

Intervention to accelerate seasonal growth. The results of this study indicate that the seasonal effect with deficit growth in the spring season was most pronounced in the middle primary grades (second and third) and were minimized for the special education population. The results of Graney et al. (2009) indicate that the seasonal effect can function so that there is deficit growth in the fall season and accelerated growth in the spring. As noted in the Introduction, the data set that was analyzed by Graney et al. (2009) derived from a university and school partnership to implement a data-based decision making and evidence-based interventions. Those factors might have influenced the patterns of seasonal effects. There is a substantial literature base to support the conclusion that data-based decision making can function to accelerate academic growth (Carnine & Granzin, 2001; Fuchs & Fuchs, 1986; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Graney et al., 2009), and there is substantial theoretical support for the use of evidence-based practices, which were both components of service delivery that were targeted by the partnership program. It might be that progress monitoring and data use accelerated learning and ameliorated the deficit growth associated with seasonal effects. Indeed, students who are served within special education have explicit goals and their progress is monitored. Although it is not possible to draw any final conclusions, it may well be that the lack of convergence in seasonal patterns between Graney et al. and other studies—along with minimal seasonal effects observed for the special education population—provides further evidence for the use of data-based decision making to improve growth during otherwise deficit periods.

Limitations and Future Research

The sample studied was entirely from five relatively ethnically homogeneous districts in a single midwestern state. Future research should attempt to corroborate these findings across a more diverse sample of students and educational programs. Finally, the results of this study derive from data collected at three points in the academic year (fall, winter, spring). Future research should examine whether the same conclusions would result from the analysis of a more robust data set (e.g., weekly, monthly) and with more precise dates that define the data collection schedule. Graney et al. (2009) correctly criticize the methods of this study and that of Ardoin and Christ (2008) because the precise dates of data collection were not used in the analysis; rather, both studies rely on data that are centered by month and not day. Although the authors of this and the prior study believe the conclusions would be substantially similar with different centering methods or more precise dates, the criticism ought to be accounted for in future research. Future research should investigate the nature of growth across summative and formative measurement conditions. Research should also examine the effect on educational decisions when the potential for seasonal effects is ignored and consistent annual growth is assumed.

Implications for School-Based Decision Making

CBM-R was developed to guide routine instructional decisions and enhance instructional effects (Deno, 1986), and it has emerged as one of the primary methods of assessment to guide problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention models of service delivery (National Association of State Directors of Education, 2005). As described in the Introduction of this study, CBM-R is used to establish expectations for both the level and rate of reading achievement throughout the primary grades; therefore, the results of this study provide at least two clear implications. First, there is good reason to question the validity of linear growth estimates as criteria to guide progress monitoring (i.e., discrepant in rate) and dual-discrepancy decisions (i.e., discrepant in both level and rate; Fuchs, Fuchs, McMaster & Al Otaiba, 2003). Second, seasonal effects can be either admired as a phenomenon or targeted for intervention. These implications are addressed in what follows.

First, if CBM-R estimates of growth are used to evaluate instructional effects and response to intervention, then those who interpret data should be aware that the magnitude of growth might be less in the spring than in the fall. There are two weighty implications that derive from this study. First, a dual-discrepancy approach to special education eligibility, which compares the observed level and rate of achievement to criterion standards (Fuchs, Fuchs, McMaster & Al Otaiba, 2003), might function to overidentify students during deficit seasons of growth and underidentify students during accelerated seasons of growth. If the data follow the pattern observed in this study, then it is likely that the dual-discrepancy model might underidentify students in the fall season and overidentify students in the spring season. That consequence is counter to the values of early intervention and prevention inherent within problem solving and response to intervention. Researchers and practitioners should pay close attention to the patterns of growth and the potential for seasonal effects to influence decisions at the local level. Accelerated rates of growth should be expected in the fall.

Second, the pattern of seasonal effects that are observed across populations and studies provide preliminary evidence that goal setting, progress monitoring and data-based decision making can minimize, eliminate or reverse the pattern of seasonal effects (e.g., Graney et al., 2009). Although future research is necessary, it is a working thesis that the systematic use of data to evaluate instructional effects and select effective instructional strategies effectively accelerates learning and prevents problems at the individual, group, and systems level. After all, data-based decision making is a central tenet of problem solving (Deno, 2005, 2002; Shinn, 2008) and response to intervention (National Association of State Directors of Education, 2005); however, data must be both systematically collected *and* systematically used to establish effects (Fuchs et al., 1992, 1991; Fuchs & Fuchs, 1986). The mere act of data collection without data use typically fails to establish effects (Fuchs & Fuchs, 1986). The explanatory hypotheses proposed in the previous section establish that deficit growth in the spring season might be remediated at the systems level if data are systematically collected and systematically used. If detected, this problem (seasonal effect) is ripe for problem analysis at the local level. Deficit growth within a single season should not be admired as an acceptable phenomenon, but, instead, it should be targeted for analysis and intervention.

Christ (2008) defined problem analysis as the "collection, summary, and use of information to systematically test, reject, or verify relevant hypotheses to establish problem solutions" (p. 159). That is, deficit rates of growth in the spring season might be remediated if expectations remain high and data are used to evaluate and maintain high standards for instructional effects. There is no reason to assume that the seasonal effect is inherent to the educational system. Instead, hypotheses related to the causal and maintaining variables should be developed—such as those presented earlier—and tested through intervention to remediate the problem. It is likely that the pattern of seasonal effects is causally related to what occurs in the classroom; and what occurs in the classroom can be influenced by the systematic used of data to evaluate of response to intervention for universal, supplemental, and intensive tiers of service delivery. This study provides evidence of seasonal effects as a potential target for intervention at the system, grade, or classroom level.

References

AIMSweb. (2002). *Reading curriculum-based measurement (R-CBM) standard oral reading fluency assessment passages.* Retrieved October 2003, from http://www.aimsweb.com

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum based measurement slope estimate using data from tri-annual universal screenings *School Psychology Review*, *37*, 109–125.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The SCRED Model: Incorporating systems-level organization and a multi-tiered problem-solving process for intervention delivery. In M. Burns, S. Jimerson, & A. VanDerHayden (Eds.), *The handbook of response to intervention: The science and practice of assessment and intervention* (pp. 321–332). New York: Springer.

Brophy, J. (1987). Synthesis of research on strategies for motivating students to learn. *Educational Leadership, 45*(2), 40–48.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2004). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43*, 61–85.

Carnine, D., & Granzin, A. (2001). Setting learning expectations for students with disabilities. *School Psychology Review*, *30,* 466–472.

Christ, T. J. (2008). Best practices in problem analysis. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 159–176). Bethesda, MD: National Association of School Psychologists.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74,* 68.

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

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15,* 358–374.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 37–56). Bethesda, MD: National Association of School Psychologists.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–42). New York: Guilford Press.

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

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). An introduction to latent variable growth curve modeling: Concepts, issues, and applications (2nd ed.). Mahwah, NY: Lawrence Erlbaum Associates.

Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling, 15,* 513–533.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*, 3–14.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10,* 43–52.

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities, 22,* 51–59.

Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 131–152). New York: Guildford Press.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436–450.

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

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research and Practice, 6,* 66–74.

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, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 431–449). New York: Guilford Press.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. *Educational Leadership, 38*(5), 415–422.

Good, T., Cooper, H., & Blakey, X. (1980). Classroom interaction as a function of teacher expectations, student sex and time of year. *Journal of Educational Psychology, 72,* 378–385.

Graney, S. B., Missall, K. N., Martinez, R. S., & 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.

Howe, K. B., & Shinn, M. M. (2002). Standard *reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Retrieved October 2004, from http://www.aimsweb.com

Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education for Students Placed at Risk, 9,* 169–188.

McKown, C., Weinstein, R. S. (2007). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*, 235–261.

National Association of State Directors of Education. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: Author.

National Reading Panel. (2000). *Report of the National Reading Panel Subgroups: Teaching children to read* (No. NIH 00–4754). Bethesda, MD: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports, 19,* 115–118.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*, 83–90.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education, 34,* 164–172.

Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision-making. *School Psychology Review, 17,* 61–80.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guildford Press.

Shinn, M. (2008). Best practices for using curriculum based measurement in a problem solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Stage, S. A. (2001). Program evaluation using hierarchical linear modeling with curriculum-based measurement reading probes. *School Psychology Quarterly, 16,* 91–112.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41,* 85–120.

Date Received: May 30, 2007
Date Accepted: April 21, 2010
Action Editor: Tanya Eckert ■

Theodore J. Christ, PhD, is an associate professor in the Educational Psychology Department at the University of Minnesota in Minneapolis—Saint Paul. He is the codirector of the Research Institute for Problem Solving. He is engaged in research and training to evaluate, develop, and enhance the tools and procedures that enable data-based decision making and problem solving within school settings. Interests include progress monitoring, curriculum-based measurement, direct behavior ratings, and computer adaptive testing.

Benjamin Silberglitt completed his PhD in Educational Psychology at the University of Minnesota at Minneapolis—Saint Paul. His areas of interest include student assessment school organization and system-level change, problem solving, response to intervention, and data-based decision making. He is director of Software Applications for TIES, a technology cooperative of 39 Minnesota school districts.

Seungsoo Yeo, PhD, is a researcher at Korean Educational Development Institute. He received his PhD from the University of Minnesota at Twin Cities in 2008. His primary research interests include curriculum-based measurement, response to intervention, and learning disabilities. His methodological work has focused on latent growth modeling and multilevel meta-analysis.

Damien C. Cormier, MA, is a school psychology doctoral student in the Educational Psychology Department at the University of Minnesota. He is a graduate research assistant at the National Center on Educational Outcomes. His research interests are in the areas of large-scale assessment policy and practice, measurement issues in individual assessment, high-functioning autism, and cultural and linguistic biases in assessment.