School Psychology Review
2018, Volume 47, No. 1, pp. 83–94
DOI: 10.17105/SPR-2017-0009.V47-1

Progress Monitoring in Reading: Comparison of Weekly, Bimonthly, and Monthly Assessments for Students at Risk for Reading Difficulties in Grades 2–4

Stacy-Ann A. January
University of South Carolina

Ethan R. Van Norman
Georgia State University

Theodore J. Christ
University of Minnesota

Scott P. Ardoin
University of Georgia

Tanya L. Eckert
Syracuse University

Mary Jane White
University of Minnesota

Abstract. The present study examined the utility of two progress monitoring assessment schedules (bimonthly and monthly) as alternatives to monitoring once weekly with curriculum-based measurement in reading (CBM-R). General education students ($N = 93$) in Grades 2–4 who were at risk for reading difficulties but not yet receiving special education services had their progress monitored via three assessment schedules across 1 academic year. Four mixed-factorial analyses of variance tested the effect of progress monitoring schedule (weekly, bimonthly, monthly), grade (2, 3, and 4), and the interaction effect between schedule and grade on four progress monitoring outcomes: intercept, slope, standard error of the estimate, and standard error of the slope. Results indicated that (a) progress monitoring schedule significantly predicted each outcome, (b) grade predicted each progress monitoring outcome except the standard error of the slope, and (c) the effect of schedule on each outcome did not depend on students' grade levels. Overall, findings from this study reveal that collecting CBM-R data less frequently than weekly may be a viable option for educators monitoring the progress of students in Grades 2–4 who are at risk for reading difficulties.

Data-based decision making is a key component of effective multitiered systems of support (MTSS). Within MTSS, all students are screened three times per year. Students who do not meet norm-referenced expectations are identified as being "at risk," and have their progress monitored frequently. Scores from frequent progress monitoring are used to make decisions about the effectiveness of core instruction and/or supplemental intervention for individual students. Despite a nearly 40-year history, research investigating the interpretations and use of scores from curriculum-based

Author Note. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A090038 to the University of Minnesota. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Correspondence regarding this article may be addressed to: Stacy-Ann A. January, Department of Psychology, University of South Carolina, 1512 Pendleton St., Columbia, SC, 29208. e-mail: sajanuary@sc.edu

measurement in reading (CBM-R) for making progress monitoring decisions at the individual level is less than robust, and many practices are based upon expert opinion or anecdotal evidence (Ardoin, Christ, Morena, Cormier, & Klingbeil, 2013; Gersten et al., 2008). For example, the impact of different progress monitoring assessment schedules for accurately estimating student growth is understood poorly. The current study addressed this important limitation in the literature by investigating differences in estimates of reading growth (as measured by CBM-R) when progress monitoring data were collected weekly, bimonthly, or monthly across the academic year for a sample of general education students in Grades 2–4.

Curriculum-Based Measurement in Reading

CBM-R is a 1 min task in which students read a grade-level passage aloud while the examiner records their errors, directly measuring students' oral reading rate with accuracy (Deno, 1985). CBM-R was originally developed to monitor students' progress in their special education curriculum (Deno, 1985, 2003); however, its use in schools has expanded tremendously. This expansion is due in part to increased emphasis in elementary schools on prevention and data-based decision making and to federal legislation (i.e., the Individuals with Disabilities Education Improvement Act of 2004) allowing for the consideration of students' responsiveness to instruction and intervention when making decisions about special education eligibility. In addition to being quick, inexpensive, and easy to administer and score, CBM-R is useful within MTSS because the resulting score is a strong estimate of students' reading achievement (Ardoin, Eckert, et al., 2013; January & Ardoin, 2015; Reschly, Busch, Betts, Deno, & Long, 2009) and can accurately distinguish among students with and without reading difficulties (January, Ardoin, Christ, Eckert, & White, 2016; Kilgus, Methe, Maggin, & Tomasula, 2014). After screening all students within a school, those identified as at risk for reading difficulties are provided with supplemental intervention and have their progress monitored frequently (e.g., once weekly; Deno et al., 2009; Silberglitt, Parker, & Muyskens, 2016). Evidence suggests that, at the group level, CBM-R is sensitive to students' growth in reading over time (Ardoin, Christ, et al., 2013; Deno, Fuchs, Marston, & Shin, 2001).

CBM-R for Progress Monitoring

Using CBM-R to monitor the progress of students with reading difficulties is fairly straightforward. First, educators set an ambitious goal of one to two words gained per week in oral reading rate (Deno et al., 2001; Fuchs, Fuchs, Hamlett, Waltz, & Germann, 1993) while students receive evidence-based intervention(s). As CBM-R data are collected, scores (which are reported in the metric of the number of words read correctly per minute [WRCM]) are plotted in a time-series fashion. Then the observed level and rate of growth in WRCM (i.e., CBM-R slope) is compared with the expected level and rate of growth (Silberglitt et al., 2016). If it is determined that the student is making adequate progress, the intervention is continued; however, if the student is not making adequate progress, the intervention may be changed, or more intensive intervention may be provided. Ultimately, students not responsive to evidence-based interventions are evaluated for special education supports. Therefore, the technical adequacy of scores from CBM-R as reflections of students' actual rate of growth—as opposed to measurement error—is one of the essential components to making accurate progress monitoring decisions (Fuchs, 2004).

When evaluating the technical adequacy of CBM-R slope (i.e., rate of growth) estimates, educators should consider two important statistics: the standard error of the estimate (*SEE)* and the standard error of the slope (*SEb*). *SEE* refers to variation around a point estimate on the line of best fit through CBM-R scores. *SEE* values are considered a reflection of the quality of the CBM-R passages, with prior research classifying *SEE* of approximately 5 WRCM as very good, 10 WRCM as good, and 15 WRCM or above as poor (Christ, Zopluoglu, Long, & Monaghen, 2012). CBM-R passage sets vary in quality, with *SEE* values reported in prior research ranging from 7 to 16 WRCM (Ardoin, Christ, et al., 2013). *SEb* refers to the variation around the slope of CBM-R scores, such that larger values indicate less precision. Imprecise estimates of growth increase the probability of making an inaccurate decision about intervention effectiveness, which ultimately may have implications for high-stakes decisions, such as special education eligibility. Together, the *SEE* and *SEb* associated with scores from CBM-R directly impact the accuracy of the decisions that are made when monitoring students' progress.

CBM-R Progress Monitoring: Impact of Schedule

Apart from universal screening three times per year, the most frequent assessment schedule for monitoring the progress of at-risk students is weekly (Mellard, McKnight, & Woods, 2009). Thus, it is not surprising that much of the progress monitoring research has focused on assessment schedules that examine the impact of multiple occasions per week (Christ, 2006; Thornblad & Christ, 2014) or one occasion per week (Ardoin, Christ, et al., 2013; Christ, Zopluoglu, et al., 2012). However, collecting, graphing, analyzing, and interpreting CBM-R data each week for every student who is having their progress monitored can be time consuming for educators (Wesson, Fuchs, Tindal, Mirkin, & Deno, 1986). Less frequent progress monitoring schedules that yield accurate data about a student's growth are desirable. Unfortunately, only a few studies have examined the utility of progress monitoring that occurs less frequently than weekly.

Using simulated data, Christ, Zopluoglu, Monaghen, and Van Norman (2013) examined the precision of slope estimates from monthly progress monitoring schedules over the course of 20 weeks. Findings revealed that reliable and valid estimates of growth for making low-stakes decisions were possible when monitoring once monthly, but only after 2–3 months using a very good (*SEE* = 5) passage set and more than 4 months with a good (*SEE* = 10) passage set. Similar results were observed in a related simulation study when data were collected using a pre–post schedule (Christ, Monaghen, Zopluoglu, & Van Norman, 2012). Despite the encouraging findings, the simulated nature of the data used in the studies limits their external validity. One of the assumptions of the data sets in these studies was that CBM-R growth equaled 1.5 WRCM per week for all students. Although this is an ambitious goal, it is likely that data collected within school settings would yield slopes that are more or less steep than 1.5 WRCM per week. Further, Christ et al. (2013) did not investigate progress monitoring schedules with frequencies that fell between weekly and monthly.

Other considerations associated with the aforementioned CBM-R simulation studies (e.g., Christ, Monaghen, et al., 2012; Christ et al., 2013) warrant further discussion. When simulating hypothetical intercepts and slopes for students, previous findings that growth magnitude depends on a student's initial level of performance (Silberglitt & Hintze, 2007) were not modeled. Additionally, although prior research (e.g., Fuchs, Fuchs, Tindal, & Deno, 1986; Fuchs, Fuchs, & Deno, 1985) demonstrated that students with lower rates of improvement have lower levels of *SEE* than students receiving instruction in general education (and presumably students exhibiting higher rates of improvement), these findings were not accounted for during data generation. Another point overlooked is that many recommendations for progress monitoring practices from the simulation studies were based upon levels of measurement error associated with ambitious values that are not regularly observed in school settings (Christ, Zopluoglu, et al., 2012). Finally, parameters for all simulations were based upon one extant data set, and despite recommendations from the authors to replicate the research using different data sets to derive parameters or modeling assumptions (Christ, Zopluoglu, et al., 2012; Christ et al., 2013), other researchers have yet to do so.

In two studies with students in Grades 2–8 eligible for special education services, Jenkins and colleagues (Jenkins, Graff, & Miglioretti, 2009; Jenkins & Terjeson, 2011) investigated a variety of monitoring schedules over an 8- or 10-week period. Jenkins et al. examined differences in slope for three schedules (every 2, 3, and 4 weeks) as compared with students' slopes from weekly progress monitoring for 10 weeks. Findings revealed that the slope for the every-3-weeks schedule had the strongest correlation with the slope from all 10 weeks of data. Jenkins et al. concluded that monitoring every 3 weeks may be the best alternative to monitoring weekly. In a related study, Jenkins and Terjeson found that the slopes for each schedule investigated (weekly, every 2, 4, and 8 weeks) were all highly related and not significantly different from each other. Together, findings from these two studies indicate that monitoring students' progress less often than weekly is promising. However, generalization of the findings from these studies is limited for at least two reasons. First, both studies included students who received special education services. This is problematic because research indicates growth rates for students in special education are different from those in general education (e.g., Christ, Silberglitt, Yeo, & Cormier, 2010; Deno et al., 2001; Graney, Missall, Martinez, & Bergstrom, 2009). Second, data were collected for only 8 or 10 weeks. Evidence suggests that at least 10–14 weeks of weekly progress monitoring data are necessary to obtain reliable and valid estimates of growth with a good-quality passage set (Christ, Zopluoglu, et al., 2012; Christ et al., 2013). Further, with only 2 months of data collection, the monthly progress monitoring schedule likely was not yet reliable and valid (Christ et al., 2013). Studies that collect progress monitoring data for a longer period may better inform decisions about alternative progress monitoring schedules.

More recently, Mercer and Keller-Margulis (2015) examined CBM-R growth by comparing data collected during universal screening (September, January, May) with data collected each month between screenings (October, November, December, February, March, April) in a sample of first- and second-grade students. The findings indicated the monthly schedule resulted in larger estimates of growth than slopes obtained with data from the two adjacent screening periods. Moreover, CBM-R slopes for the two schedules had a moderate, statistically significant association in the spring of first grade but were not significantly related for the students in second grade. As such, the use of a monthly monitoring schedule was not recommended. However, one important limitation was that the study did not include estimates of weekly growth, and therefore the extent to which the monthly slopes reflected students' weekly growth in WRCM was unknown.

CBM-R Progress Monitoring: Impact of Grade

In contrast to research on alternative progress monitoring schedules, there is a larger body of literature supporting differences in level and slope of CBM-R scores as a function of grade. Differences in level are most apparent when examining the universal screening benchmark goals for various grades; WRCM benchmark scores increase with grade level. The opposite pattern is evident when examining growth in CBM-R scores. That is, gains in oral reading rate across time are greater for younger students when collecting universal screening (Christ et al., 2010; Nese et al., 2012) and progress monitoring data (Deno et al., 2001; Fuchs et al., 1993). For students receiving general education services, published research suggests weekly growth decreases across grade levels, with estimates of gains in WRCM being approximately 1.5, 1, and .85–.95 in second, third, and fourth grades, respectively (Deno et al., 2001; Fuchs et al., 1993; Silberglitt & Hintze, 2007). Despite evidence for differences in CBM-R slopes as a function of grade level, no studies have investigated whether the effect of grade on slope magnitude depends on the type of progress monitoring schedule utilized. Knowing whether the impact of progress monitoring schedule varies for students in different grades will help practitioners make important decisions about how often to monitor students' progress.

 DOI: 10.17105/SPR-2017-0009.V47-1

Current Study

The most common CBM-R progress monitoring schedules used in research and practice are triannually for universal screening and weekly progress monitoring assessments for students who are identified as at risk (Mellard et al., 2009). However, few studies have examined the utility of bimonthly or monthly progress monitoring schedules, which could substantially reduce the time and resources needed to monitor students' response to intervention. The research that does exist is limited, as studies did not include a variety of progress monitoring schedules (e.g., Mercer & Keller-Margulis, 2015), included students receiving special education services (e.g., Jenkins et al., 2009), or collected data for a relatively short period of time (e.g., Jenkins & Terjeson, 2011). Furthermore, extant research relied on simulated progress monitoring data (e.g., Christ et al., 2013) or did not account for student grade level (e.g., Jenkins et al., 2009). Therefore, the extent to which findings from extant research are appropriate for informing progress monitoring decisions among general education students being monitored in school settings across the academic year is unclear.

The current study builds upon and extends existing research on the frequency of progress monitoring assessment schedules by comparing intercept, slope, *SEE*, and *SEb* for three schedules (weekly, bimonthly, monthly) with general education students identified as at risk who were monitored across the entire academic year. Moreover, this study collected data in elementary school settings and simultaneously evaluated the impact of grade. This study sought to investigate three research questions. Given the body of research indicating that progress monitoring outcomes differ as a function of grade (Deno et al., 2001; Fuchs et al., 1993), we were primarily interested in Research Question 1 (the impact of assessment schedule on outcomes), and Research Question 3 (whether the impact of schedule depended on grade level).

1. What is the effect of schedule (weekly, bimonthly, monthly) on progress monitoring outcomes (intercept, slope, *SEE*, *SEb*)?
2. What is the effect of grade (2, 3, 4) on progress monitoring outcomes (intercept, slope, *SEE*, *SEb*)?
3. Does the effect of schedule on progress monitoring outcomes differ as a function of grade?

METHOD

Schools were recruited initially as part of a large, multiyear study aimed to develop and validate CBM-R probes for screening and progress monitoring in elementary school. The schools that participated in this study varied in their MTSS practices. Prior to the project, most schools engaged in some combination of universal screening, progress monitoring, and the implementation of academic interventions with students identified as at risk. The current study includes data from 1 year of the project, with participating schools located in the Midwest, Southeast, and Northeast regions of the U.S.

Participants

Participants included 93 students enrolled in Grades 2 ($n = 32$), 3 ($n = 27$), and 4 ($n = 34$) in one of five schools, 45% ($n = 42$) of whom were female. Participants were 67% ($n = 62$) White, 22% ($n = 20$) Hispanic or Latino, 8% ($n = 7$) Black or African American, and 3% ($n = 3$) American Indian, Asian, or Pacific Islander. Students were included in the current study if their median CBM-R score from the fall universal screening fell below the 40th percentile (based on national norms) and they did not receive special education services. We elected to include only students at risk for academic difficulties but not yet receiving special education services for two reasons. First, progress monitoring data collected from students receiving special education services often differ in statistically significant ways relative to data collected from students not receiving special education services (Deno et al., 2001; Fuchs et al., 1986; Fuchs et al., 1985). Given the sample size, it would be difficult to statistically control for special education status in analyses. Second, we elected to exclude students who were not identified as at risk to promote the external validity of findings. Students not at risk for later difficulties are often monitored only during triannual universal screening periods (Silberglitt et al., 2016). Thus, by collecting weekly progress monitoring data only on students with at least some risk in reading, the findings from this study are more generalizable than if progress monitoring data were collected on students who would not typically be monitored frequently. Institutional review board approval, parental consent, and student assent were obtained prior to data collection.

Measure

CBM-R passages from FastBridge Learning (Christ et al., 2014), a publisher of universal screening and progress monitoring assessments, were used in the current study. Additional measures from another publisher were administered as part of the larger project but are not relevant to the current study. The administration order of each measure was randomized and counterbalanced for each student at each weekly data collection occasion. The decision to use only FastBridge passages in this study was made for two reasons. First, differences in progress monitoring outcomes as a function of passage type have been addressed elsewhere (see Ardoin & Christ, 2009). Second, the primary purpose of this study was to evaluate differences in progress monitoring outcomes as a function of assessment schedule; only the FastBridge passages were administered via the assessment schedules described herein.

In each grade, 20 progress monitoring FastBridge probes are available. Test–retest, alternate forms, and interrater reliability estimates for FastBridge passages range from .75 to .94 across Grades 2–4 (Christ et al., 2014). Concurrent validity estimates with the Test of Silent Reading Efficiency and Comprehension range from .79 to .81. Similarly, concurrent validity coefficients for static scores with aimsweb and

DIBELS Next probes range from .95 to .97 and from .92 to .96, respectively. Multilevel reliability of growth estimates ranges from .45 to .86, and split-half reliability estimates range from .90 to .98 across Grades 2–4.

Experimenter Training

Experimenters were retired teachers and other individuals who had previous experience working with children, undergraduate research assistants, and graduate students. Prior to independently conducting an assessment, researchers with expertise in CBM-R first trained experimenters using exercises and videos until they demonstrated 100% reliability across three practice probes. Then, experimenters watched an CBM-R administration session conducted in the field, were supervised during their first administration, and were provided with performance feedback.

Procedures

Beginning in September, students were monitored weekly across the entire school year. Each week, students were administered one FastBridge CBM-R probe. Note that there were not separate data collection points that specifically corresponded with a bimonthly schedule (a description of how the bimonthly observations were analyzed is given later). Once per month, students were administered three more CBM-R probes, in addition to the one FastBridge passage. These three probes were different from the passages administered weekly and were the same across the duration of the study. Efforts were made to administer measures on the same day each week; on occasions when this was not possible, students were administered the probe(s) on the nearest day. Due to minor differences in school calendars, the average total number of weeks student progress was monitored was 33 (*SD* = 2 weeks). Further, data could not be collected when school was not in session (e.g., holiday breaks). When administering any probe, experimenters read a scripted protocol to each participant and followed standardized FastBridge administration procedures (Christ et al., 2014). All data collection took place in separate, quiet areas within the schools, including general education classrooms, vacated classrooms or offices, and hallways. Across all students and measurement occasions, 57.15% of observations were collected in a hallway, 31.38% in an empty office, 4.98% in an empty classroom, 0.26% in the back of a classroom during regular instruction, and 6.23% in other settings. The variety of data collection settings reflect typical in-school testing locations and researchers accommodating school staff requests to not disrupt daily schedules.

Overall assessment fidelity was measured by the primary investigators using an instrument created for the project and exceeded 97% on average. This instrument (available from the third author) was adapted from the Accuracy of Implementation Rating Scale used within the AIMSweb system. Project personnel used the instrument to monitor the accuracy in which data collectors presented stimuli (e.g., placed correct copy of probe in front of student), read directions (e.g., read exact examiner prompts as written for each probe type), timed (e.g., started and stopped timer correctly), and scored passages (e.g., marked errors correctly).

Analytic Plan

Several steps were taken to facilitate data analyses. For progress monitoring outcomes associated with a weekly schedule, all available weekly data were analyzed and were regarded as the best estimate of students' actual growth. Each student was expected to have 30 weeks of CBM-R data. Across an average of 33 weeks of data collection, each student had 30 weekly progress monitoring assessments; thus, no data were missing. For the bimonthly data collection schedule, every other observation from the weekly data collection schedule was analyzed, whereas for the monthly data collection schedule, the average WRCM of all three probes was used in analyses. Next, time was coded as the number of weeks since the first observation for each schedule. As such, all slope values can be interpreted as the average increase in WRCM per week, and progress monitoring outcomes are comparable across data collection schedules. Observations from either the weekly, bimonthly, or monthly data collection schedules were used to estimate each trend line. The dependent variables (i.e., progress monitoring outcomes) in this study were intercept, slope, *SEE*, and *SEb*, with a particular interest in the latter three outcomes.

Primary data analyses occurred in several steps. First, three ordinary least squares (OLS) trend lines (one corresponding with each data collection schedule) were estimated for each student. Although slope estimates from OLS regression have potential shortcomings for estimating growth for individual students (Mercer, Lyons, Johnston, & Millhoff, 2015; Vannest, Parker, Davis, Soares, & Smith, 2012), it was chosen in this study for two reasons: (a) it is the most recommended procedure for estimating slopes for individual students (e.g., Ardoin, Christ, et al., 2013; Christ, Zopluoglu, et al., 2012; Good & Shinn, 1990), and (b) it is the predominant approach being used by practitioners in schools, via several major publishers of CBM probes. Next, the values of each progress monitoring outcome (i.e., intercept, slope, *SEE*, and *SEb*) were extracted from each schedule for each student. Descriptive statistics were calculated for intercept, slope, *SEE*, and *SEb* values, disaggregated by grade and data collection schedule. Third, a series of separate mixed-factorial analyses of variance (ANOVAs) were conducted with each progress monitoring outcome as a dependent variable. Grade served as the between-subjects factor, and data collection schedule was the within-subjects factor in each analysis. Finally, we conducted planned comparisons using paired-samples *t* tests based upon the findings of the series of ANOVAs. For the ANOVAs, alpha was set to .004 using a Bonferroni correction (.05/12) to control for Type I error as three terms (main effect for grade, main effect for schedule, and the

 DOI: 10.17105/SPR-2017-0009.V47-1

interaction of grade and schedule) were evaluated across four models (intercept, slope, *SEE*, and *SEb* as outcomes). Critical *p* values for the planned contrasts were set based upon the number of comparisons made in the ANOVAs. Although Cohen's (1988) rules of thumb for interpreting effect sizes (≤.29 = small, .30–.50 = medium, and >.50 = large) are not without limitations (Lakens, 2013; Richardson, 2011), they were used in this study to contextualize statistically significant results.

RESULTS

Descriptive statistics for each grade by progress monitoring schedule are presented in Table 1. Across data collection schedules, intercept and slope values displayed a similar pattern. Intercept values tended to increase as grade level increased, and slope values tended to decrease as grade level increased. For instance, when considering a weekly data collection schedule, the average intercept values for students in Grades 2, 3, and 4 were equal to 47.71 (*SD* = 20.11), 59.17 (*SD* = 14.92), and 83.97 (*SD* = 19.11) WRCM, respectively. Considering only monthly data collection schedules, the average slopes for Grades 2, 3, and 4 were 1.37 (*SD* = 0.34), 1.15 (*SD* = 0.56), and 1.10 (*SD* = 0.47) WRCM of improvement per week, respectively. Regarding the influence of schedule on intercept and slope values, intercept values tended to increase across grades as the frequency of data collection increased. The opposite pattern was observed for slope values.

Across data collection schedules, *SEE* values tended to increase as a function of grade level (see Table 1). Considering only the bimonthly data collection schedule, average *SEE* values were equal to 8.05 (*SD* = 1.76), 8.61 (*SD* = 2.18), and 8.97 (*SD* = 2.01) in Grades 2, 3, and 4, respectively. A similar yet less pronounced pattern was observed with *SEb* values. Considering data from weekly data collection schedules, average *SEb* values were 0.15 (*SD* = 0.04), 0.17 (*SD* = 0.03), and 0.18 (*SD* = 0.04) in Grades 2, 3, and 4, respectively. Results for *SEE* and *SEb* as a function of schedule were less consistent. For instance, average *SEE* values increased as the frequency of data collection schedules increased in Grade 2 but not in Grade 3.

Effect of Schedule and Grade Level on Progress Monitoring Outcomes

In all ANOVAs, the Greenhouse–Geisser adjustment was applied due to a violation of sphericity (Stevens, 1996), which resulted in degrees of freedom values with decimal points. Significant main effects were observed for grade and progress monitoring schedule for intercept, $F(2, 90) = 35.49$, $p < .001$; $F(1.38, 124.20) = 45.98$, $p < .001$, and slope, $F(2, 90) = 15.50$, $p < .001$; $F(1.30, 117) = 103.28$, $p < .001$, values, respectively (see Table 2). Grade level and data collection schedule constituted a medium effect (η_p^2 = .44 and .34, respectively) in explaining variability in intercept values. For slope, grade constituted a small to medium effect (η_p^2 = .26), and data collection schedule had a medium to large effect

Table 1. Descriptive Statistics for Progress Monitoring Outcomes

Grade	Outcome	Monthly Progress Monitoring				Bimonthly Progress Monitoring				Weekly Progress Monitoring			
		M	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
2 (*n* = 32)	Intercept	42.65	17.71	1.69	3.60	45.98	20.90	1.57	3.07	47.71	20.11	1.55	2.67
	Slope	1.37	0.34	–0.77	0.89	1.12	0.35	0.89	2.10	1.02	0.30	–0.14	0.82
	SEE	7.72	2.06	0.82	0.08	8.05	1.76	0.29	–0.45	8.43	1.81	1.26	2.72
	SEb	0.15	0.05	0.79	–0.14	0.20	0.05	0.24	0.04	0.15	0.04	1.13	2.42
3 (*n* = 27)	Intercept	53.26	13.79	0.54	–0.30	58.07	14.33	0.26	–0.56	59.17	14.92	0.52	–0.23
	Slope	1.15	0.56	0.73	0.64	0.80	0.42	0.89	0.18	0.72	0.43	1.29	1.61
	SEE	8.01	1.74	0.36	0.01	8.61	2.18	–0.27	–1.00	9.18	1.78	–0.13	–1.23
	SEb	0.15	0.03	0.00	–0.31	0.22	0.05	–0.25	–0.97	0.17	0.03	0.02	–1.39
4 (*n* = 34)	Intercept	78.21	18.80	0.69	–0.96	81.89	19.00	0.58	–1.07	83.97	19.11	0.52	–1.08
	Slope	1.10	0.47	0.52	0.47	0.59	0.23	–0.29	0.13	0.48	0.25	–1.06	1.01
	SEE	9.71	1.91	1.01	1.60	8.97	2.01	0.03	–1.21	9.65	1.88	–0.07	–1.12
	SEb	0.18	0.04	1.15	2.03	0.23	0.05	0.17	–1.03	0.18	0.04	0.06	–1.17

Note. Growth estimates from the monthly progress monitoring assessment schedule were based upon the average of three words read correctly per minute (WRCM) scores collected from three curriculum-based measurements in reading passages once a month. Growth estimates from the bimonthly progress monitoring assessment schedule were based upon one WRCM score collected once every other week. Growth estimates from the weekly progress monitoring assessment schedule were based upon one WRCM score collected once every week.

Table 2. Results From Mixed-Factorial Analyses of Variance for Each Progress Monitoring Outcome

Source (Type)	SS	*df*	MS	*F*	*p*	η_p^2
Intercept						
Grade (B)	66674.70	2	33337.35	35.49	<.001	.44
Grade (E)	84530.37	90	939.23			
Schedule (W)*	1512.76	1.38	756.38	45.98	<.001	.34
Grade x Schedule (I)*	20.35	2.76	5.09	0.31	.818	.01
Grade x Schedule (E)*	2960.77	124.20	16.45			
Slope						
Grade (B)	9.97	2	4.98	15.50	<.001	.26
Grade (E)	28.95	90	0.32			
Schedule (W)*	11.14	1.30	5.57	103.28	<.001	.53
Grade x Schedule (I)*	0.78	2.60	0.20	3.62	.029	.07
Grade x Schedule (E)*	9.71	117	0.05			
Standard Error of the Estimate						
Grade (B)	95.42	2	47.71	6.28	.003	.12
Grade (E)	683.68	90	7.60			
Schedule (W)*	20.64	1.50	10.32	6.16	.003	.06
Grade x Schedule (I)*	19.21	3	4.80	2.87	.038	.06
Grade x Schedule (E)*	301.45	135	1.67			
Standard Error of the Slope						
Grade (B)	0.03	2	0.02	4.09	.020	.08
Grade (E)	0.37	90	<0.01			
Schedule (W)*	0.18	1.60	0.09	112.51	<.001	.56
Grade x Schedule (I)*	0.01	3.20	<0.01	1.91	.131	.07
Grade x Schedule (E)*	0.14	144	<0.01			

Note. Critical *p* values were set at .004 to control Type I errors due to interpreting effects from conducting multiple analyses of variance. SS = sum of squares; MS = mean square; η_p^2 = partial eta squared; B = between-subjects factor; E = error; W = within-subjects factor; I = interaction term.
*Adjustment made to degrees of freedom using Greenhouse–Geisser ε for violation of sphericity to test the statistical significance of results.

(η_p^2 = .56). The grade by schedule interaction term was not statistically significant for either ANOVA. Grade level and schedule accounted for a statistically significant amount of variation in *SEE* values, $F(2, 90) = 6.29$, $p = .003$; $F(1.50, 135) = 6.16$, $p = .003$ (see Table 2). However, the practical significance of each predictor was negligible (η_p^2 = .12 and .06, respectively). Similar to the intercept and slope results, the interaction term for grade and schedule was not statistically significant, $F(3, 135) = 2.87$, $p = .038$. For *SEb*, the main effect for schedule was statistically, $F(1.60, 144) = 112.51$, $p < .001$, and practically (η_p^2 = .56) significant. The effect of grade level, $F(2, 90) = 4.09$, $p = .020$, and the interaction term for grade level and schedule $F(3.20, 144) = 1.91$, $p = .131$, were not statistically significant.

Planned Comparisons

Given that the primary focus of this study was to investigate the influence of data collection schedules on progress monitoring outcomes and the lack of statistically significant interaction terms for any progress monitoring outcome, all planned comparisons were focused on the main effect of schedule. Thus, we conducted a series of paired-samples *t* tests to evaluate the influence of the main effect of schedule on progress monitoring outcomes, collapsed across grades. To control for Type I errors, we applied a Bonferroni correction (.05/12 = .004) to assess the statistical significance of results (see Table 3).

Table 3 presents results from the planned comparisons. For intercept values, all contrasts were statistically significant.

Table 3. Results for Planned Comparisons for Progress Monitoring Outcomes as a Function of Assessment Schedule

				Paired Samples *t* Test					
		Marginal Means		Bimonthly			Weekly		
Outcome	Schedule	*M*	*SE*	R – C	95% CI	*p*	R – C	95% CI	*p*
Intercept	Monthly	58.73	6.09	3.89	2.46, 5.32	<.001	5.56	4.29, 6.82	<.001
	Bimonthly	62.62	6.49	–	–	–	1.67	0.97, 2.37	<.001
	Weekly	64.29	6.67	–	–	–	–	–	–
Slope	Monthly	1.21	0.13	–0.38	–0.46 , –0.28	<.001	–0.47	–0.55, –0.40	<.001
	Bimonthly	0.83	0.09	–	–	–	–0.09	–0.13, –0.06	<.001
	Weekly	0.74	0.09	–	–	–	–	–	–
SEE	Monthly	8.53	0.88	0.02	–0.43, 0.47	0.936	0.56	0.14, 0.99	.001
	Bimonthly	8.55	0.89	–	–	–	0.54	0.30, 0.79	<.001
	Weekly	9.10	0.89	–	–	–	–	–	–
SEb	Monthly	0.16	0.02	0.06	0.05, 0.07	<.001	0.01	0.00, 0.01	.211
	Bimonthly	0.22	0.02	–	–	–	–0.05	–0.06, –0.04	<.001
	Weekly	0.17	0.02	–	–	–	–	–	–

Note. Critical *p* values for paired samples *t* tests were set at .004 to control for Type I errors for multiple contrasts (.05/12). For paired samples *t* tests, results are for mean of differences between rows (R) and columns (C). Growth estimates from the monthly progress monitoring assessment schedule were based upon the average of three words read correctly per minute (WRCM) scores collected from three curriculum-based measurements in reading passages once a month. Growth estimates from the bimonthly progress monitoring assessment schedule were based upon one WRCM score collected once every other week. Growth estimates from the weekly progress monitoring assessment schedule were based upon one WRCM score collected once every week. *SEE* = standard error of the estimate; *SEb* = standard error of the slope.

The marginal mean for intercepts was largest when data were collected weekly (*MM* = 64.29, *SE* = 6.67) and smallest when data were collected monthly (*MM* = 58.73, *SE* = 6.09). As with intercepts, all contrasts for slope values were statistically significant. However, marginal means for slope values decreased as the frequency of data collection increased for monthly (*MM* = 1.21, *SE* = 0.13), bimonthly (*MM* = 0.83, *SE* = 0.09), and weekly (*MM* = 0.74, *SE* = 0.09) progress monitoring schedules. There was a statistically significant difference between weekly (*MM* = 9.10, *SE* = 0.89) and monthly (*MM* = 8.53, *SE* = 0.88; $t[92] = 2.64$, $p = .001$) *SEE* values. There was also a statistically significant difference in *SEE* values when data were collected weekly versus bimonthly (*MM* = 8.55, *SE* = 0.89; $t[92] = 4.40$, $p < .001$). For *SEb* values, a statistically significant difference was observed between outcomes based upon bimonthly (*MM* = 0.22, *SE* = 0.02) and monthly (*MM* = 0.16, *SE* = 0.02; $t[92] = 11.13$, $p < .001$) data collection schedules, as well as between weekly (*MM* = 0.17, *SE* = 0.02) and bimonthly data collection schedules ($t[92] = -15.93$, $p < .001$).

DISCUSSION

As part of a data-based decision-making model such as MTSS, students who are at risk for significant reading difficulties have their progress monitored frequently, and resultant scores are used to make decisions about the effectiveness of instruction or intervention. Often, educators choose to monitor students' progress weekly (Mellard et al., 2009); however, the utility of less frequent progress monitoring schedules for accurately estimating growth among students at risk for later difficulties is unclear. If less frequent schedules are viable, educators can use less time and fewer resources monitoring students' progress and dedicate more time to providing effective instruction to remediate academic skill deficits. The current study extended prior research by collecting progress monitoring data from general education students who were at risk for reading difficulties (but were not receiving special education services) in authentic school settings across the entire academic year. We investigated the utility of monitoring progress weekly, bimonthly, or monthly by testing the effects of schedule (Research Question 1), grade (Research Question 2), and the interaction between schedule and grade (Research Question 3) on four progress monitoring outcomes: intercept, slope, *SEE*, and *SEb*. Given that the latter three outcomes were of greatest interest, we will focus most of the discussion on slope, *SEE*, and *SEb*.

Regarding the effect of progress monitoring schedule on outcomes, findings indicated that the type of progress monitoring schedule was a significant predictor of intercept, slope, *SEE*, and *SEb*. Less frequent progress monitoring schedules were associated with larger estimates of slope. These findings suggest that monitoring students' reading less frequently than weekly may reduce the accuracy of estimates of students' growth (i.e., slope). The bimonthly schedule yielded a slope estimate that was most similar to the slope estimate observed with the weekly schedule (within .09 WRCM per week). In contrast, the slope from the monthly progress monitoring schedule overestimated growth by approximately 0.5 words per week. The fact that slope estimates for the bimonthly and monthly schedules were larger than the slope for the weekly schedule is important, given the reliance on slope when interpreting progress monitoring data. In contrast, the monthly and bimonthly progress monitoring schedules resulted in the lowest *SEE* values. However, the difference among *SEE* values was relatively small and may reflect the consistent quality level of the CBM-R passage set (Christ et al., 2013). Finally, of the three monitoring schedules, the *SEb* associated with bimonthly monitoring was significantly greater than the *SEb* for the monthly and weekly schedules. Notably, one possible explanation for the similar *SEb* values between the monthly and weekly schedules could be related to the fact that the average of three scores (as opposed to a single score) was used for the monthly schedule. Although the difference in *SEb* values between the bimonthly and weekly schedule was small in magnitude (.05) in this study, when combined with an inaccurate slope, the practical significance may not be trivial. Conversely, previous progress monitoring research suggests that the duration of progress monitoring schedules, which was fixed for all comparisons, is a major influential factor impacting the precision of growth estimates (Christ, 2006; Christ et al., 2013).

Findings from this study add to an emerging body of research investigating the impact of assessment schedule on CBM-R progress monitoring outcomes. For example, results do not support the use of a monthly monitoring schedule, which is consistent with a field study of assessment schedules (Mercer & Keller-Margulis, 2015) but not with a simulated study (Christ et al., 2013). It is also important to note that, across grades, a weekly CBM-R slope of .74 WRCM was observed, which was approximately half the magnitude of the slopes assumed in simulation studies (e.g., Christ, Monaghen, et al., 2012; Christ, Zopluoglu, et al., 2012; Christ et al., 2013). One potential explanation is that the simulation studies represent an ideal case for growth in oral reading rate. Conversely, it is possible that the longer duration of progress monitoring in this study impacted the slope estimates (Christ, 2006; Christ et al., 2013). Finally, in comparison with previous research using students receiving special education services (Jenkins et al., 2009; Jenkins & Terjeson, 2011), current findings yielded a different pattern of results. In the two studies by Jenkins and colleagues less frequent monitoring schedules resulted in slope estimates that were smaller in magnitude as compared with the weekly schedule. The difference between current findings and prior research (Jenkins et al., 2009; Jenkins & Terjeson, 2011) was expected and could be due to a few different factors, including the fact that data were collected in prior studies over a shorter duration (8–10 weeks) and the qualitative differences between students receiving general versus special education services. Nonetheless, this inconsistency, coupled with previous research (Deno et al., 2001; Fuchs et al., 1986; Fuchs et al., 1985), further underscores the need to examine the impact of assessment schedule and progress monitoring outcomes separately for students who are receiving special education services and those who are not.

The second and third research questions addressed the main effect of grade and the interaction effect of grade and schedule, respectively. Consistent with prior research (e.g., Deno et al., 2001; Silberglitt & Hintze, 2007) and reading theory (Chall, 1983), results indicated that younger students demonstrated more growth across the academic year than older students, and initial level (intercept) of WRCM was greater for older students than for younger students. Grade level also impacted *SEE*, and based on the descriptive results, there was an increasing pattern of WRCM as grade level increased. However, this effect was in the small range. Finally, there was no significant interaction between schedule and grade for any of the progress monitoring outcomes (i.e., no moderation effects were observed). This finding is important, as it indicates that the effect of schedule on intercept, slope, *SEE*, or *SEb* did not differ as a function of grade.

Implications for Practice

Findings from this study have important implications for practitioners engaging in the progress monitoring of students not meeting reading benchmarks. First, results of this study indicate that progress monitoring with CBM-R less frequently than once per week may be a viable option. When comparing bimonthly and monthly progress monitoring, a bimonthly schedule was found to be most appropriate estimate of weekly growth in WRCM of students in Grades 2–4. This recommendation is based largely on our findings that the intercept and slope from the bimonthly schedule were closest to the intercept and slope of the weekly schedule and that differences in *SEE* and *SEb* were relatively small. Monitoring monthly is not recommended, primarily due to its overestimation of weekly growth. Moreover, present findings reveal that the differences in progress monitoring schedules did not depend on the grade level of students. This suggests that a bimonthly monitoring schedule can be used with students in second, third, and fourth grades. Notably, it is important to underscore that findings from this study do not necessarily support the use of CBM-R progress monitoring data for making high-stakes decisions, such as special education eligibility. Rather, findings from this study can be used as a starting point to help school psychologists and educators make decisions about how often to monitor the progress of students in Grades 2–4 who are being administered CBM-R.

Limitations and Future Research

Strengths of the study include that data were collected in school settings, and that general education students were monitored across an entire school year, allowing for estimates of growth for the population of students who are most likely to be monitored frequently. Thus, compared with studies that used a relatively short duration, simulated data, or included students who received special education services, the current findings likely generalize better to students who are at risk for reading difficulties. Despite the study's strengths, there are also potential limitations that should be considered. First, participants in the current study were administered CBM-R passages from one passage set (FastBridge); therefore, the extent to which the current findings generalize to passage sets from other publishers is unknown (Ardoin & Christ, 2008).

A second limitation of the current study was the relatively small size of the sample in each grade. This was due in part to the time and resources required to collect the data in this study and the inclusion criteria that students were at risk yet were still receiving general education services. Third, it is unknown which students in the current study may have been receiving additional intervention services across the academic year and when instructional changes may have occurred. Finally, given data were collected across the academic year, it is possible that, at least for some students, weekly growth was not linear (Van Norman & Parker, 2016). However, we chose to use OLS regression for the current analyses, as this is current practice in schools. Moreover, test vendors continue to produce normative growth rates based upon the assumption of linear growth within a school year. Future research should investigate the extent to which progress monitoring data collected from a bimonthly assessment schedule is linear.

Given the current findings that bimonthly progress monitoring may be an alternative to weekly progress monitoring, research is needed to investigate further the conditions under which a bimonthly schedule is most appropriate for progress monitoring. For instance, given the varied quality of passage sets, the minimum number of bimonthly sessions necessary to produce reliable and valid growth estimates is unclear. This is important in light of extant research suggesting that 12–16 weeks of monitoring may be necessary before being able to make a technically adequate progress monitoring decision (Christ et al., 2013). Similarly, given that the average duration of progress monitoring in this study was 33 weeks, it is possible that findings may differ for shorter durations of time. Thus, future research is needed to investigate the impact of the type and duration of progress monitoring schedules on outcomes for students in elementary school. This information will be important in better understanding how long to monitor students' growth before making decisions about intervention effectiveness. Future studies should also investigate the predictive validity of slopes from bimonthly schedules in estimating students' reading achievement and the decisions (e.g., changes in intervention intensity) that are made using progress monitoring data. Finally, future research may investigate the viability of alternative monitoring schedules for other progress monitoring measures.

Conclusions

This study sought to investigate the potential utility of two CBM-R assessment schedules (bimonthly, monthly) as alternatives to once weekly progress monitoring of students who are at risk for reading difficulties. Because previous research demonstrated that grade level impacts rate of growth per week (e.g., Deno et al., 2001), we were also interested in the effect of grade on progress monitoring outcomes and whether the effect of grade depended on schedule. Findings from this study indicated that schedule and grade were significant predictors of intercept, slope, and *SEE*, whereas only schedule was a predictor of *SEb*. However, none of the interaction effects between grade and schedule were statistically significant, suggesting that the effect of schedule on the outcomes did not differ as a function of grade. Follow-up contrasts for the main effect of schedule indicated that (a) the monthly schedule had the smallest intercept, (b) the bimonthly schedule had a slope that was closest to the estimate of weekly growth, (c) the bimonthly and monthly schedules yielded *SEE* values that were slightly smaller than the weekly schedule, and (d) the monthly schedule had a similar *SEb* value to the weekly schedule. Overall, the pattern of findings from this study indicate that, as compared with a monthly schedule, a bimonthly progress monitoring schedule is most accurate in estimating the weekly growth in CBM-R scores for students at risk for reading difficulties in grades 2–4.

CONFLICT OF INTEREST STATEMENT

Theodore J. Christ, PhD, has equity and royalty interests in, and will serve on the board of directors for, FastBridge Learning (FBL) LLC, a company involved in the commercialization of the Formative Assessment System for Teachers (FAST). The University of Minnesota also has equity and royalty interests in FBL LLC. These interests have been reviewed and managed by the University of Minnesota in accordance with its conflict of interest policies. None of the other authors have conflicts of interest to declare.

Scott P. Ardoin, PhD, has equity in Theodore J. Christ and Colleagues, LLC (TJCCC) which owns the CBM-R probes used in the commercialization of the Formative Assessment System for Teachers (FAST).

REFERENCES

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

Ardoin, S. P., & Christ, T. J. (2009). Curriculum-based measurement of oral reading: Standard errors associated with progress monitoring outcomes from DIBELS, AIMSweb, and an experimental passage set. *School Psychology Review*, *38*, 266–283.

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding curriculum-based measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology*, *51*, 1–18. doi:10.1016/j.jsp.2012.09.004

Ardoin, S. P., Eckert, T. L., Christ, T. J., White, M. J., Morena, L. S., January, S.-A. A., & Hine, J. F. (2013). Examining variance in reading comprehension among developing readers: Words in context (curriculum-based measurement in reading) versus words out of context (word lists). *School Psychology Review*, *42*, 243–261.

Chall, J. S. (1983). *Stages of reading development*. New York, NY: McGraw Hill.

Christ, T. J. (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*, 128–133.

Christ, T. J., Arañas, Y. A., Kember, J. M., Kiss, A. J., McCarthy-Trentman, A., Monaghen, B. D., … White, M. J. (2014). *Formative Assessment System for Teachers technical manual: earlyReading, CBMReading, aReading, aMath, and earlyMath*. Minneapolis, MN: Formative Assessment System for Teachers.

Christ, T. J., Monaghen, B. M., Zopluoglu, C., & Van Norman, E. R. (2012). Curriculum-based measurement of oral reading: Evaluation of growth estimates derived with pre–post assessment methods. *Assessment for Effective Intervention*, *38*(3), 139–153. doi:10.1177/1534508412456417

Christ, T. J., Silberglitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review*, *39*, 447–462.

Christ, T. J., Zopluoglu, C., Long, J. D., & Monaghen, B. D. (2012). Curriculum-based measurement of oral reading: Quality of progress monitoring outcomes. *Exceptional Children*, *78*, 356–373. doi:10.1177/001440291207800306

Christ, T. J., Zopluoglu, C., Monaghen, B. D., & Van Norman, E. R. (2013). Curriculum-based measurement of oral reading: Multi-study evaluation of schedule, duration, and dataset quality on progress monitoring outcomes. *Journal of School Psychology*, *51*, 19–57. doi:10.1016/j.jsp.2012.11.001

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

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219–232. doi:10.1177/001440298505200303

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*, 184–192. doi:10.1177/00224669030370030801

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.

Deno, S. L., Reschly, A. L., Lembke, E. S., Magnusson, D., Callender, S. A., Windram, H., & Stachel, N. (2009). Developing a school-wide progress-monitoring system. *Psychology in the Schools*, *46*, 44–55. doi:10.1002/pits.20353

Fuchs, D., Fuchs, L. S., Tindal, G., & Deno, S. L. (1986). Performance instability of learning disabled, emotionally handicapped, and nonhandicapped children. *Learning Disability Quarterly*, *9*(1), 84–88. doi:10.2307/1510405

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, *33*, 188–192.

Fuchs, L. S., Fuchs, D., & Deno, S. L. (1985). Importance of goal ambitiousness and goal mastery to student achievement. *Exceptional Children*, *52*, 63–71. doi:10.1177/001440298505200108

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.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to intervention (RTI) and multi-tier intervention in the primary grades* (NCEE 2009-4045). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/rti_reading_pg_021809.pdf

Good, R. H., III, & Shinn, M. R. (1990). Forecasting accuracy of slope estimates for reading curriculum-based measurement: Empirical evidence. *Behavioral Assessment*, *12*(2), 179–193.

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. doi:10.1016/j.jsp.2008.12.001

January, S.-A. A., & Ardoin, S. P. (2015). Technical adequacy and acceptability of curriculum-based measurement and the Measures of Academic Progress. *Assessment for Effective Intervention*, *41*(1), 3–15. doi:10.1177/1534508415579095

January, S.-A. A., Ardoin, S. P., Christ, T. J., Eckert, T. L., & White, M. J. (2016). Evaluating the interpretations and use of curriculum-based measurement in reading and word lists for universal screening in first and second grade. *School Psychology Review*, *45*, 310–326. doi:10.17105/SPR45-3.310-326

Jenkins, J., Graff, J. J., & Miglioretti, D. L. (2009). Estimating reading growth using intermittent CBM progress monitoring. *Exceptional Children*, *75*, 151–163. doi:10.1177/001440290907500202

Jenkins, J., & Terjeson, K. J. (2011). Monitoring reading growth: Goal setting, measurement frequency, and methods of evaluation. *Learning Disabilities Research & Practice*, *26*(1), 28–35. doi:10.1111/j.1540-5826.2010.00322.x

Kilgus, S. P., Methe, S. A., Maggin, D. M., & Tomasula, J. L. (2014). Curriculum-based measurement of oral reading (R-CBM): A diagnostic test accuracy meta-analysis of evidence supporting use in universal screening. *Journal of School Psychology*, *52*, 377–405. doi:10.1016/j.jsp.2014.06.002

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12. doi:10.3389/fpsyg.2013.00863

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice*, *24*(4), 186–195. doi:10.1111/j.1540-5826.2009.00292.x

Mercer, S. H., & Keller-Margulis, M. A. (2015). Consistency and magnitude of differences in reading curriculum-based measurement slopes in benchmark versus strategic monitoring. *Psychology in the Schools*, *52*, 316–324. doi:10.1002/pits.21823

Mercer, S. H., Lyons, A. F., Johnston, L. E., & Millhoff, C. L. (2015). Robust regression for slope estimation in curriculum-based measurement progress monitoring. *Assessment for Effective Intervention*, *40*, 176–183. doi:10.1177/1534508414555705

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*(4), 887–915. doi:10.1007/s11145-011-9304-0

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*, 427–469. doi:10.1016/j.jsp.2009.07.001

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. doi:10.1016/j.edurev.2010.12.001

Silberglitt, B., & Hintze, J. M. (2007). How much growth can we expect? A conditional analysis of R-CBM growth rates by level of performance. *Exceptional Children*, *74*, 71–84. doi:10.1177/001440290707400104

Silberglitt, B., Parker, D., & Muyskens, P. (2016). Assessment: Periodic assessment to monitor progress. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden, (Eds.), *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (2nd ed., pp. 271–291). New York, NY: Springer.

Stevens, J. P. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.

Thornblad, S. C., & Christ, T. J. (2014). Curriculum-based measurement of reading: Is 6 weeks of daily progress monitoring enough? *School Psychology Review*, *43*, 19–29.

Van Norman, E. R., & Parker, D. C. (2016). An evaluation of the linearity of curriculum-based measurement of oral reading (CBM-R) progress monitoring data: Idiographic considerations. *Learning Disabilities Research & Practice*, *31*(4), 199–207. doi:10.1111/ldrp.12108

Vannest, K. J., Parker, R. I., Davis, J. L., Soares, D. A., & Smith, S. L. (2012). The Theil–Sen slope for high-stakes decisions from progress monitoring. *Behavioral Disorders*, *37*, 271–280.

Wesson, C., Fuchs, L., Tindal, E., Mirkin, P., & Deno, S. L. (1986). Facilitating the efficiency of on-going curriculum-based measurement. *Teacher Education and Special Education*, *9*(4), 166–172.

Date Received: February 7, 2017
Date Accepted: January 12, 2018
Associate Editor: Kelli Cummings ■

AUTHOR BIOGRAPHICAL STATEMENTS

Stacy-Ann A. January, PhD, NCSP, is an assistant professor in the Department of Psychology at the University of South Carolina. Dr. January's research interests include (a) investigating the technical characteristics and decision-making utility of universal screening and progress monitoring assessments and (b) evaluating interventions that target academic or behavioral skills.

Ethan R. Van Norman, PhD, is an assistant professor in the school psychology program in the Department of Counseling and Psychological Services at Georgia State University. His research primarily focuses on assessment practices within tiered systems of support. In addition, Van Norman conducts research aimed at building the capacity of educators and school psychologists to use data meaningfully to make sound educational decisions.

Theodore J. Christ, PhD, is a professor of school psychology in the Department of Educational Psychology and co-director of the Research Institute for Problem Solving (RIPS) at the University of Minnesota. It was his work in those roles that established him as the founder and chief scientific officer of FastBridge Learning (fastbridge.org). Dr. Christ is engaged to innovate techniques and technology to serve professional educators who seek to improve educational outcomes. He is interested in and engaged with the content, methodology, analytics, technology, software, and people who pursue the same.

Scott P. Ardoin, PhD, BCBA, is codirector of the Center for Autism and Behavioral Education Research and department head of the Department of Educational Psychology at the University of Georgia. His recent work employs eye-tracking technology to (a) measure changes in the reading behavior of students as a function of interventions and (b) examine the extent to which test and question format alter students' reading behavior and test-taking strategies. Dr. Ardoin's research is published in numerous refereed journals. He is a recipient of the APA Division 16 Lightner Witmer Award, a recipient of the UGA Creative Research Medal, and currently president-elect of the Society for the Study of School Psychology.

Tanya L. Eckert, PhD, is an associate professor of psychology at Syracuse University. Her research interests include examining procedures for assessing academic and behavioral problems, developing classroom-based interventions, and measuring the acceptability of assessment and intervention procedures. Dr. Eckert is senior associate editor of *School Psychology Review.*

Mary Jane White, PhD, is a research associate at the University of Minnesota. She earned her degree in educational psychology with an emphasis in reading comprehension. She is currently project manager for Dr. Theodore Christ's federally funded grant work and supports bridging his and other colleagues' research into practice through FastBridge Learning.