Technical Report # 1701

Oral Reading Fluency:
Outcomes from 30 Years of Research

Gerald Tindal

University of Oregon

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
http://brt.uoregon.edu

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

Abstract

This paper is about oral reading fluency, how it has been measured, the students who have been measured, and the outcomes that have been reported for both performance and progress. This measurement system has almost become ubiquitous with documentation of response to intervention in reading and become a focal point of research on reading. Fluency is one of the 'Big Five' skill areas identified by the National Reading Panel and provides an accessible target for both researchers and practitioners. In the course of 30 years, oral reading fluency has been administered in the field in a relatively standard manner, though the reading passages and populations associated with its use, as well as the manner for analyzing results, have varied. In the end, the results are remarkably consistent, both in levels of performance and rates of progress.

Key Words: *Reading fluency, reading measurement, curriculum-based measurement of reading*

Oral reading fluency: Outcomes from 30 Years of Research

Research on growth in oral reading fluency[1] has been conducted for over 30 years and in this time the methodology has reflected interesting gaps as well as curious consistencies. From a methodological vantage point, one of the most important shifts has been in the population studied over this time with increasing attention to student subpopulations. Instrumentation has also shifted from informal, researcher crafted measures (typically using random samples of passages from grade level texts) to more standardized (and published) instruments. Finally, the manner in which growth has been analyzed has drastically changed with sophisticated multilevel and latent growth models more prevalent in recent work.

Introduction to the Measurement of Oral Reading Fluency

The focus on oral reading fluency began in the mid 1970s with publication of *Data Based Program Modification: A Manual* (S. L. Deno & Mirkin, 1977). This monograph launched the eventual development and validation of curriculum-based measurement (CBM). The premise behind this measurement system was to democratize and localize a collection of high-quality student performance and progress measures so that teachers could evaluate the effects of their instruction. Though the manual primarily articulated a problem-solving model for program development in special education, it was based on the use of behavioral measures that would be useful for classroom teachers.

These measures emphasized standardized administration in reading, writing, and math that would take little time to learn and implement, be capable of frequent use, and provide information that was reliable and led to valid decisions. Soon after this publication, the *Institute for Research on Learning Disabilities* (IRLD) was funded at the University of Minnesota with

[1] Oral reading fluency is typically assessed using text passages though a few studies have used word lists; in this manuscript, the focus is exclusively on students orally reading from text passages (not word lists) under timed conditions (one minute) to establish the number of words read correctly.

the charge to investigate then current learning disabilities identification systems and provide alternatives for teachers to use. One of the first measures to be investigated in reading was oral reading fluency (ORF), which was generally defined as the number of words read correctly in one minute. Conventions for administration and scoring were studied and validated; in the end, ORF became one of the most heavily studied and widely adopted measures of reading and was specifically operationalized as the number words read in total (in one minute) from a randomly sampled passage of text minus the number of words read incorrectly (misidentified, hesitated in reading, or omitted), resulting in words ready correctly per minute (*wcpm*). Over the course of five years, a number of technical reports were published on both traditional and alternative identification systems; for CBM, these reports addressed both technical adequacy and a number of issues on adoption and use.

In the initial report published on the technical adequacy of ORF, reliability and criterion-referenced validity was documented (Deno, Mirkin, & Chiang, 1982; Deno, Mirkin, Chiang, & Lowry, 1980). Further studies were conducted to document the utility of oral reading fluency including the following.

1. Growth of oral reading fluency (both in reading from word lists and from passages) over the year (from fall to winter to spring) for students in grades one to six (Marston, Lowry, Deno, & Mirkin, 1981).

2. Effect of varying item domains (grade specific or across grades) and sample duration (30 and 60 seconds) (Fuchs, Tindal, & Deno, 1981).

3. Improvement of instructional practices with better academic gains, more structured instruction, more realistic goals, and more accuracy in judging student performance and progress (Fuchs, Deno, & Mirkin, 1982).

4. Application of CBMs in a norm-referenced manner in which students' standing in a group may provide an appropriate reference for growth and program evaluation (Marston & Magnusson, 1985; Tindal, Germann, & Deno, 1983; Tindal, Germann, Marston, & Deno, 1983).

As this research was accumulating and confirming the utility of CBM in general and ORF specifically, its application was becoming more widespread with other researchers adopting practices of varying similarity. In 1985, a special issue of *Exceptional Children* was devoted to curriculum-based assessment, of which curriculum-based measurement was but one variety (Tucker, 1985). Three important books were published, two of them consolidating and summarizing the research basis of the measurement system (Shinn, 1989, 1998) and the other expanding the measurement system to focus on multiple references – norm, criterion, and individual (Tindal & Marston, 1990). Two practical guides to CBM have been published more recently by M. K. Hosp, Hosp, and Howell (2007) and J. L. Hosp, Hosp, Howell, and Allison (2014). Finally, and most recently, (see Tindal, 2013) conducted an extensive summary of CBM.

Research on fluency moved out of the CBM platform with its inclusion in the National Assessment of Educational Progress (NAEP), often considered the nation's report card. In this study, three dimensions of reading were measured: accuracy (conformance to letter-sound conventions), rate (speed of reading aloud), and fluency, a "distinct attribute of oral reading separate from accuracy and rate. Fluency was defined in terms of phrasing, adherence to the author's syntax, and expressiveness and was measured at one of four levels" (National Assessment of Educational Progress, 2002, p. v). A major conclusion from this study was that "the three separate oral reading abilities— accuracy, rate, and fluency— are related to each other, and all three are related to reading comprehension" (p. v). Perhaps most importantly, 44% of tested students were found to be dysfluent.

Nearly 20 years passed between the initial research validation studies and the major findings from the National Reading Panel (2000) with the inclusion of fluency as one of the 'big five' components for teaching children to read (the other four being phonemic awareness, phonics, vocabulary, and text comprehension). Fluency was defined as "the ability to read a text quickly, accurately, and with proper expression" (p. 3-5) and invoked the concept of 'automaticity', a concept first identified as having an important relation with comprehension (LaBerge & Samuels, 1974). The main findings from the NRP review of the literature included the following.

1. Guided repeated oral reading (with feedback) improved reading; the average effect size from meta-analyses of this practice was 0.41.
2. Independent reading was an important component of becoming a fluent reader though the results were much less encouraging than repeated reading.
3. There was "a close relationship between fluency and reading comprehension. Students who are low in fluency may have difficulty getting the meaning of what they read" (p. 3-1).
4. Fluency measurement can be conducted with informal reading inventories, miscue analysis, or reading speed calculations but all of them require "oral reading of text, and all can be used to provide an adequate index of fluency" (3-9).
5. The reason for these findings is likely due to greater fluency freeing additional reading 'cognitive resources' and thus allowing the reader to group words into meaningful grammatical units for interpretation.

We began our literature search with two recent summaries of technical adequacy for curriculum-based measures of reading: (a) Wayman, Wallace, Wiley, Tich'a, and Espin (2007) and (b) Ardoin, Christ, and Morena (2013). Both reviews were systematic and thorough,

comprehensively covering considerable literature on curriculum-based reading measures, though both were far broader than our current interest in growth of oral reading fluency. For example, Wayman et al. (2007) addressed the student sample (count, grades, and level), the reading measures used (maze and word reading fluency), time and scoring procedures, and finally the main findings for validity, reliability, and growth or slope. Though they included reliability and validity data (which is beyond our scope in this paper), they also considered slope (which is exactly what we address in this paper). The Ardoin, Christ, and Morena review was primarily focused on decision rules when using oral reading fluency, which provided a list of studies for which time series data (growth) had been collected. We also expanded our literature search to include earlier studies in the investigation of ORF growth (beginning with the initial systematic research that had begun with the Institute of Research on Learning Disabilities at the University of Minnesota under the leadership of Deno). Finally, we covered more current studies that have been published since the Wayman et al. research and updated a few studies on growth published since the Ardoin, Christ, and Morena publication by including research not limited to decision-making.

Populations (Samples and Demographics)

The initial research published on growth of oral reading fluency scores was conducted to document the technical adequacy for a measurement system using data based program modification (the primary purpose for which curriculum-based measures were developed). Initial research on oral reading fluency included relatively few students either sampled around Minneapolis, where the research institute was located, or Pine County, where district wide implementation of CBM was adopted. The first three publications were technical reports (Marston et al., 1981; Tindal, Germann, & Deno, 1983; Tindal, Germann, Marston, et al., 1983),

with the last one sampling students in an effort to document normative performance with well over 200 students in each grade level. The last study, which was published in the professional literature, included 64 students in Grades 3 to 5 (Fuchs, Deno, & Mirkin, 1984). Students participating in a data based program modification treatment were compared to the performance of 77 students with teachers who set goals and monitored progress as they wished. All students were labeled with a handicap (the term used at that time for students receiving special education services). However, once the technical adequacy of the measures was established and initial 'norms' reported for district wide use, a significant amount of research was devoted to better understanding typical growth in ORF, thereby serving as an impetus to measure fluency for a large number of students. Another early ORF study was conducted by Marston and Magnusson (1985) who compared students in general and special education (Grades 1 to 6) at three time periods (fall, winter, and spring) in a large urban school district. Almost a decade later, Fuchs, Fuchs, Hamlett, Walz, and Germann (1993) measured over 3,000 students in Grades 1-6 but only 374 in reading with the specific aim of establishing typical slopes of improvement for ORF.

Student sample size and breadth. The critical issue in sampling a large number of students in each grade level has been to ensure results are generalizable. However, it wasn't until the publication by Hasbrouck and Tindal (1992) that the sample of students began to be sufficiently large and broadly represented (geographically) that they could be considered as 'normative'. This publication was designed to specifically reflect representative values because of the diverse range of regions and curriculum sampled, though no count of students per grade was presented. A later and more detailed publication by the same authors (Hasbrouck & Tindal, 2006) reported norms with several thousand students in each grade at each of three periods (fall, winter, and spring). Again, the sample included a wide range of students from districts

throughout the country that had adopted ORF measurement using various curricula or assessment tools. Though studies continue to be published with small sample sizes, often to study special populations, researchers have continued to document changes in ORF with increasingly large student samples.

Small sample studies. A number of studies have measured ORF with *hundreds of students per grade level*. MacMillan (2000) used a Rasch model to analyze growth in ORF for nearly 1,700 students in Grades 2-7. Shortly thereafter, growth norms were published using a regionally stratified sample of 3,000 students in Grades 1-6 (S. L. Deno, Fuchs, Marston, & Shin, 2001). Both Silberglitt and Hintze (2005) and Puranik, Petscher, Al Otaiba, Catts, and Lonigan (2008) studied nearly 2,000 students from first to third grade, with the latter study confined to speech and language impaired students. Graney, Missall, Martinez, and Bergstrom (2009) studied growth for almost 900 students in Grades 3-5 over two academic years, with students from a single elementary school in the rural Midwest. This same type of longitudinal design was used by Wanzek et al. (2010), with the oral reading fluency of over 400 students tracked from Grade 1 through 3 to predict state and norm-referenced tests. Almost 5,000 students' ORF in Grades 2-6 were measured by Christ, Silberglitt, Yeo, and Cormier (2010) in an effort to document growth rates and seasonal effects. Yeo, Fearrington, and Christ (2011) measured oral reading fluency of more than 1,700 students in Grades 3-8 to document gender, income, and special education status 'bias' on curriculum-based measurement slope in reading (though the term bias was never explicated) and later studied the relation between oral reading fluency and maze performance with over 1,500 students (Yeo, Fearrington, & Christ, 2012). Similarly, nearly 2,500 students were measured on ORF to document within-year changes by Nese et al. (2012) and later included nearly 1,500 students across Grades 1-8 (with eight measurement occasions)

to study the functional form of growth (Nese et al., 2013). Finally, Tolar, Barth, Fletcher, Francis, and Vaughn (2014) assessed over 1,300 students in grades 6-8 to investigate groups of readers and treatments, as well as administration conditions.

A number of studies have included *well over 1,000 students per grade level*. In the earliest published study (Marston & Magnusson, 1985), oral reading fluency measures were administered to 1,200 – 1,500 students per grade level so they could integrate decision-making (screening, identification, program planning, progress monitoring, and program evaluation). In an effort to establish growth norms, Silberglitt and Hintze (2007) measured more than 7,500 students in Grades 2-6 (all from Minnesota) while Christ and Silberglitt (2007) measured ORF on 8,200 students in Grades 1-5. The ORF performance for a large number of students was measured by Baker et al. (2008), with nearly 5,000 students in each of Grades 1-3 (all of whom were participating in Reading First schools). In a similar effort to monitor growth of a special population, Al Otaiba et al. (2009) sampled a couple of thousand students in each of Grades 2 and 3, with a sampling plan limited to Latino students. In the study by Kim, Petscher, Schatschneider, and Foorman (2010), over 13,000 students were measured in Grades 1-3, again targeting a specific population (students in Reading First schools). Wang, Algozzine, Ma, and Porfeli (2011) measured over 5,000 second graders in a large urban public school system in North Carolina. Keller-Margulis, Clemens, Im, Kwok, and Booth (2012) as well as Kamata, Nese, Patarapichayatham, and Lai (2013) had around 2,000 students per grade (3-5) in their sampling plans.

Three studies have had *substantial sample sizes of students*, with over 23,000 students in grade 1 (Schatschneider, Wagner, & Crawford, 2008), nearly 10,000 students sampled in each of

three grades (1-3) in a study conducted by Crowe, Connor, and Petscher (2009), and over 56,000 students in each of Grades 1 to 3 (Logan & Petscher, 2010).

As a final note, the students for whom growth has been documented have been primarily in Grades 1-5 with only a few studies addressing students in Grades 6-8 (Jenkins, Graff, & Miglioretti, 2009; Nese et al., 2013; Tolar et al., 2014; Yeo et al., 2012). No studies of growth in oral reading fluency have been published with students in Grades 9-12.

Specific populations. Between the extremes of studies reporting growth on scores and hundreds of students sampled (per grade) versus tens of thousands students in the sample in each grade, are a number of studies where growth is documented by a single specific characteristic of the population.

1. The initial research by Fuchs et al. (1984) focused on '*handicapped'* students and a decade later, the same distinction was made in student sampling with *'non handicapped'* students as reported by Fuchs et al. (1993). For Marston and Magnusson (1985) as well as S. L. Deno et al. (2001) and Graney et al. (2009), general and special education students were sampled.
2. Students with *specific learning disabilities and other disabilities* have been targeted by a number of researchers (D. Fuchs, Fuchs, Mathes, & Simmons, 1997; Jenkins et al., 2009; Jenkins & Terjeson, 2011; Jenkins, Zumeta, & Dupree, 2005; Simmons, Fuchs, Fuchs, Mathes, & Hodge, 1995; Tindal, Flick, & Cole, 1992). Puranik et al. (2008) reported monthly growth for *speech and language impaired students* (that was further disaggregated by persistent versus resolved or changed to a learning disability). *Struggling* versus *typical* students were compared in the most recent study by Tolar et al. (2014) with the general conclusion that ORF predicted performance on the criterion measure for typical but not struggling students (using familiar rather than novel reading passage administrations).

3. *English language learners* were the focus for several researchers (Al Otaiba et al., 2009; D. L. Baker, Park, & Baker, 2012; S. K. Baker & Good, 1995; Keller-Margulis et al., 2012; Ramírez & Shapiro, 2007).

4. As noted above, students in *Reading First schools* (with its emphasis on students of poverty and with low reading performance) were the target population for a few large scale studies (S. K. Baker et al., 2008; Schatschneider et al., 2008); similarly, the target has been populations with *lower SES* (eligible for free or reduced price lunch) versus *non-low SES* students (Crowe et al., 2009) or at risk due to poverty and/or language (Logan & Petscher, 2010).

5. *Other specific subpopulations* have been targeted in documenting ORF performance and progress. In most of this research, results are not disaggregated and reported for separate subpopulations with the following exceptions. *Low, middle, and high scoring* first grade students were studied to document differences in oral reading growth (Plascenia-Peinado, Deno, & Johnson, 2005). Speece and Ritchey (2005) as well as Logan and Petscher (2010) included *risk of failing to learn* as a defining characteristic in defining their populations. Yeo et al. (2011) described growth rates for students in grades three to eight, conditional on *free and reduced price lunch, gender, and special education status*. Finally, Nese et al. (2012) focused on within-year growth as a function of specific student groups (*student gender, free and reduced-price lunch, ethnicity, special education status, and limited English proficiency*).

In some of these studies, the targeted subpopulations were used to structure the analysis and report comparative performance. Fuchs et al. (1984) and Fuchs et al. (1997) compared pre and post performance on ORF for *experimental* versus *contrast* conditions (basically with and without data-based program modification); for Marr, Algozzine, Nicholson, and Dugan (2011) the treatment was *coaching* versus *control*. Finally, the effect of growth on struggling students

who received an *intervention* were compared to struggling students who *did not receive an intervention*, though the specifics of the intervention were not described (Tolar et al., 2014).

Baker and Good (1995) compared *English only* students with *bilingual students*; likewise, *non-English language learners, English language learners, and former English language learners* in Grades 3-5 were compared in the study by Keller-Margulis et al. (2012). Both S. L. Deno et al. (2001) as well as Graney et al. (2009) compared students in *general* versus *special* education. Students with *specific disabilities* were disaggregated in the results reported by Puranik et al. (2008) and by Al Otaiba et al. (2009). For Crowe et al. (2009), the results were disaggregated by *socioeconomic status*. Speece and Ritchey (2005) as well as Logan and Petscher (2010) compared students *at risk* of not learning to read versus those *not at risk* with the latter study adding poverty and language as risk factors. Finally, Silberglitt and Hintze (2007) divided the population into *deciles* for reporting growth (with all students from districts in Minnesota).

Measurement Issues and Treatment Sensitivity

A number of studies report growth as a function of a characteristic of the measurement system with less regard to the student population being studied or controlled by a specific student characteristic with no cross-population effects documented.

1. Both Tindal et al. (1992) and Hintze and Shapiro (1997) compared different types of reading programs (literature based versus traditional). Christ and Silberglitt (2007) focused on standard error of measurement but also reported fall, winter, and spring oral reading fluency scores for several thousand students in Grades 1-5. Ardoin and Christ (2008) investigated growth as a function of probe sets (and use of single versus multiple screening administrations). In a similar manner, Ardoin and Christ (2009) documented difference in ORF growth as a

function of reading passages (DIBELS, AIMSweb, and an experimental passage); population characteristics were known for the group only (and included only race and free and reduced price lunch). Likewise, the effect of passages was documented by Riley-Heller, Kelly-Vance, and Shriver (2005), noting differences between curriculum dependent passages and a curriculum independent passage (Test of Oral Reading Fluency). Most recently, familiar (same passages repeatedly administered) versus novel (new passages repeatedly administered) were compared with some evidence that when predictor and criterion measures were aligned; "progress monitoring slope has a stronger effect in the novel condition than in the familiar condition" (Tolar et al., 2014, p. 55).

2. Two studies focused on establishing a relation between growth in oral reading fluency and other measures such as math or reading maze (Graney et al., 2009; Yeo et al., 2012). Only student gender and subsidized lunch were reported for the group as a whole for the former study, with added information on race and special education status also provided in the latter study.

3. In a study by Jenkins et al. (2009), the focus was on growth with various schedules of progress monitoring with 41 students in Grades 3-8. Later, Jenkins and Terjeson (2011) investigated reading growth as a function of measurement frequency and methods of evaluation. In both studies, only students with disabilities were included.

4. Wood (2006) sampled close to 100 students in Grades 3-5 to better document the effect of classroom influences on growth of oral reading fluency (using hierarchical linear models). In a similar study but with school level effects, growth rates were reported as a function of general and special education status and student gender with a second grade sample (Wang et al., 2011).

5. Finally, Kamata et al. (2013) interpreted oral reading proficiency from a large population of students in Grade 4 (with no known demographics) for the express purpose of applying a growth mixture modeling approach for within-year non-linear trends using only three data points.

6. Four studies have been reported on the growth rates as a function of specific interventions. Two studies by the Fuchs team reported the effects of experimental teaching and tutoring (Simmons et al., 1995) and peer-assisted learning (Simmons et al., 1995). Marr et al. (2011) also compared treatment versus control conditions in gains made in ORF from fall to winter to spring, specifically examining an intervention designed to build fluency through coaching with special and general education students disaggregated. In all three studies, the main comparison was by the independent variable and no other specific student information included. Though no specific intervention was specified in the study by Tolar et al. (2014), no variability was found in the slope of progress monitoring for struggling students receiving no intervention when measured with novel passages every two months.

Instrumentation and Analyses of ORF Data Sets

Two issues are addressed methodologically that address the manner in which oral reading fluency is documented to establish growth rates or standards. The first critical issue is the materials (passages) being used to document changes in performance over time. The second issue is the time intervals within the year and the manner in which data are analyzed, both of which are related (e.g., non-linear growth is not possible to document with only two time points).

Measures used in documenting growth. Research on oral reading fluency began with passages typically drawn from a variety of curricula and thus reflecting the original intentions of the original authors of curriculum-based measurement (Deno & Mirkin, 1977) serve as general

outcome measures. The initial research conducted on oral reading fluency likewise reflected this broad kind of passage sampling technique Deno, Mirkin, & Chiang (1982). Throughout the 1980s and 1990s (through 2006), researchers used various curricula that were either *unspecified* (S. K. Baker & Good, 1995; S. L. Deno et al., 2001; Dunn & Eckert, 2002; L. S. Fuchs et al., 1993; Hasbrouck & Tindal, 1992, 2006; Hintze, Daly, & Shapiro, 1998; MacMillan, 2000; Speece & Ritchey, 2005; Tindal, Germann, & Deno, 1983; Tindal, Germann, Marston, et al., 1983) or *specified*, including the following specific reading curricula: Ginn 720 (Fuchs et al. 1984; Marston & Magnusson, 1985); Silver Burdett and Ginn (Stage & Jacobsen, 2001), Scribner and SRI materials (Tindal et al., 1992), Vanderbilt materials using folktales (D. Fuchs et al., 1997; Jenkins et al., 2009; Simmons et al., 1995), Edcheckup (Jenkins et al., 2005), Scott-Foresman with the Test of Reading Fluency (Riley-Heller et al., 2005) and Oral Reading Fluency CBM-Passage Fluency (ORF-PF) (Francis, Barth, Cirino, Reed, & Fletcher, 2008) or CBM-R (Silberglitt & Hintze, 2005). Otherwise, most recent studies have used standard passages from publically available instruments, primarily DIBELS, AIMSweb, and easyCBM.

Researchers using *DIBELS* passages in their studies include those associated with the Florida Center for Reading Research and various Florida institutions of higher education (IHEs) with access to statewide databases. For example, a number of studies report results using *Florida Progress Monitoring and Reporting Network [PMRN]* database or Reading First risk assessments (Al Otaiba et al., 2008; Al Otaiba et al., 2009; Crowe et al., 2009; Kim et al., 2010; Logan & Petscher, 2010; Puranik et al., 2008; Schatschneider et al., 2008) or use data associated with the Center on Teaching and Learning (CTL) at the University of Oregon (S. K. Baker et al., 2008). A recent UO CTL research study used DIBELS Next (Cummings, Park, & Bauer-Shaper, 2013). Other, independent researchers have also studied growth with DIBELS (Ardoin & Christ,

2008; D. L. Baker et al., 2012; Jenkins & Terjeson, 2011; Marr et al., 2011; Wang et al., 2011; Wood, 2006).

AIMSweb also has been extensively studied to establish growth rates or standards. Generally, the research team using AIMSweb are associated with Christ, Silberglitt, or Ardoin and draw data from statewide databases (Ardoin & Christ, 2009; Christ & Silberglitt, 2007; Christ et al., 2010; Silberglitt & Hintze, 2007; Yeo et al., 2011; Yeo et al., 2012). Other researchers using AIMSweb include Graney et al. (2009); Keller-Margulis et al. (2012); Ramírez and Shapiro (2007).

Finally, three studies have been published on growth with *easyCBM* (Kamata et al., 2013; Nese et al., 2012) with the most recent publication by Nese, Biancarosa, Cummings, Kennedy, Alonzo, and Tindal (2013).

Frequency of measurement. Ironically, though oral reading fluency was designed for frequent use as part of progress monitoring to evaluate instructional programs, most research has used pre-post or benchmark seasonal measurement intended to screen all students (as opposed to monitoring those at risk for poor reading outcomes). It is relatively rare that growth of ORF is documented using pre and post measures: Fuchs et al. (1984), Fuchs et al. (1997), Simmons et al. (1995), and Crowe et al. (2009) administered measures to document specific treatment effects. Rather, the vast majority of studies have administered the measures at benchmark times (fall, winter, and spring) that generally take place in September-October, December-January, and April-May, respectively. Only a few studies have actually used progress measures more than the three or four times that comprise benchmark or seasonal measures:

- Tindal et al. (1992) administered ORFs nearly twice per week for more than 7 months.
- Fuchs et al. (1993) administered 7+ measures.

- Hintze, Shapiro, and Lutz (1994) obtained ORF data twice per week for 9 weeks with average of 17 administrations.
- Plascenia-Peinado, Deno, and Johnson (2005) administered Grade 1 passages every week for six weeks.
- Fuchs et al. (1997) had 15 weeks of ORF to document effects from students in two conditions: Peer Assisted Learning (PALS) or control.
- MacMillan (2000) used six reading probes at each grade and three norming periods in October, November, and December, with one passage per administration.
- Hintze et al. (1998), conducted progress-monitoring sessions twice a week during a 10-week period with each student having a maximum of 20 sessions.
- Dunn and Eckert (2002) administered 32 passages read over eight weeks.
- Speece and Ritchey (2005) collected ORFs over 20 weeks (weekly for 6 weeks and then monthly from January to May).
- Riley-Heller et al. (2005) had three passages administered twice each week for 5 weeks.
- Jenkins et al. (2009) administered a total of 29 measures in 1, 2, 3, 4 weeks and pre-post (Sept.-Nov.).
- Ardoin and Christ (2009) collected ORF data over 12 weeks (with approximately 2 administrations per week) with time of year unknown.
- Jenkins and Terjeson (2011) administered measures every 2, 4, and 8 weeks.
- Nese, Biancarosa, Cummings, Kennedy, Alonzo, and Tindal (2013) analyzed data that had been collected in Grade 1 with 6 administrations and in Grades 2-8 with 8 administrations delivered to all students, not just those at-risk.

- Finally, Tolar et al. (2014) administered oral reading fluency measures every two months for a total of five administrations.

Analysis for documenting growth. Most studies documented growth simply using raw gain score or linear regression (ordinary least squares) over two or three time points. Because of this limited number of data points, analyses were necessarily restricted to linear models of growth and did not incorporate non-linear models, though Kamata et al. (2013) documented this possibility using structural equation models.

In the initial research, growth was calculated using raw score differences and converting to weekly gain by either multiplying by 5 or 7 (L. S. Fuchs et al., 1982; Marston et al., 1981; Tindal et al., 1992; Tindal, Germann, & Deno, 1983; Tindal, Germann, Marston, et al., 1983). The two large sample presentations of ORF growth by (Hasbrouck & Tindal, 1992) and the more recent publication by (Hasbrouck & Tindal, 2006) also were based on simple raw score gains. Likewise, raw scores were used to document treatment or program effects (D. Fuchs et al., 1997; Graney et al., 2009; Marston & Magnusson, 1985; Ramírez & Shapiro, 2007; Simmons et al., 1995) as well as measurement effects (Ardoin & Christ, 2008; Christ & Silberglitt, 2007; Jenkins et al., 2005).

As ORF was gaining traction in its use in schools past the initial documentation of technical adequacy, ordinary least squares (OLS) regression was being used to calculate slopes (typically over three time periods) for assessing growth in practical applications for documenting effects from curriculum and measurement variables (Ardoin & Christ, 2009; S. K. Baker & Good, 1995; S. L. Deno et al., 2001; Dunn & Eckert, 2002; L. S. Fuchs et al., 1993; Hintze et al., 1998; Hintze et al., 1994; Jenkins et al., 2009; Jenkins & Terjeson, 2011; Marr et al., 2011; Riley-Heller et al., 2005).

An increasing number of researchers have been using more advanced multilevel models or latent growth models to better understand growth at level-1 (over time) as a function of either student (or teacher) characteristics at levels-2 or -3 (Al Otaiba et al., 2009; D. L. Baker et al., 2012; S. K. Baker et al., 2008; Christ et al., 2010; Crowe et al., 2009; Keller-Margulis et al., 2012; Kim et al., 2010; Logan & Petscher, 2010; Nese et al., 2012; Puranik et al., 2008; Schatschneider et al., 2008; Silberglitt & Hintze, 2007; Speece & Ritchey, 2005; Stage & Jacobsen, 2001; Wang et al., 2011; Wood, 2006; Yeo et al., 2011; Yeo et al., 2012). In addition, latent growth modeling has been used (Nese et al., 2013; Tolar et al., 2014) to study the functional form of growth (e.g., linear, quadratic, cubic) or other variables (administration conditions, population invariance, etc.). This growth modeling has also included building probabilistic predictions of performance on state and norm-referenced tests (Wanzek et al., 2010).

Outcomes on ORF Performance and Change over Time

The vast majority of studies reveal a number of consistencies in both the levels of performance at various time periods and the change over time for oral reading fluency. For example, one of the earliest studies and a recent study show nearly the same levels of performance at all time periods. For Grades 3-6, the levels in all three time periods (fall, winter, and spring) for students reported by Tindal, Germann, and Deno (1983) were very close to those reported by (Yeo et al., 2012), even though the two studies are separated by thirty years, sample vastly different students, and use different measures.

Using the median values reported Hasbrouck and Tindal (1992, 2006) as an anchor, it appears Grade 2 students begin the year reading about 50 wcpm and improve to 90 wcpm, which represents an increase of more than one word per week. Students in Grade 3 begin the fall

reading about 70-80 wcpm and improve at a rate of 1 word per week (to about 110+ in the spring). Grade 4 students begin in the fall reading just less than 100 wcpm and improve to about 120 wcpm (reflecting a growth of just less than 1 word per week). For 5th grade students, fall oral reading fleuncy is just over 100 wcpm and improves to about 130, again showing a gain of about one word per week.

In comparison with other researchers who also measure students at known (seasonal) time periods in multiple grades, these values are fairly similar to the results from Tindal, Germann, and Deno (1983) in Grades 3-6; from Marston and Magnusson (1985) in Grades 1-6; Deno, Fuchs, Marston, and Shin (2001) in Grades 2-6; Silberglitt and Hintze (2005) and Christ and Silberglitt (2007) in Grades 2-3 and 2-5, respectively; Graney, Missall, Martinez, and Bergstrom (2009) in Grades 3-5 (with lower values for students receiving special education); Kim, Petscher, Schatschneider, and Foorman (2010) in Grades 2-3; Keller-Margulis, Clemens, Im, Kwok, and Booth (2012) in Grades 3-5; Nese, Biancarosa, Anderson, Lai, Alonzo, and Tindal (2012) in Grades 3 and 4 (with slightly higher values in grade 5); Yeo et al. (2012) in Grades 3-7; and finally, Crowe, Connor, and Petscher (2009) in Grades 2 and 3.

Not all results have been consistent, however, with the growth reported by Hasbrouck and Tindal (1999, 2006). The results from MacMillan (2000) showed higher levels in Grade 3 and lower values (with almost no differences) in Grades 4-7. Similarly, the values were slightly lower for Ramirez and Shapiro (2006) in Grades 3-5 as well as for Wanzek et al. (2010) in Grades 1-3; Baker, Smolkowski, Katz, Fien, Seeley, Kame'enui, and Beck (2008) and Baker, Park, and Baker (2012) reported lower values Grades 2 and 3; Christ, Silberglitt, Yeo, and Cormier, (2010) reported slightly higher values in Grades 3-5. The values reported by Wood (2006) are higher in Grades 3-5.

For studies in which only a single grade was measured in ORF, the results vary inconsistently from the values reported by Hasbrouck and Tindal (1992; 2006). Schatschneider, Wagner, and Crawford (2008) report similar values for Grade 1 students; Hintze, Shapiro, and Lutz (1994) report higher values for Grade 3 students; similarly, Baker and Good (1995) report higher values for Grade 2 students; Stage and Jacobsen (2001) report higher ORF for Grade 4 students. In two studies with 2nd grade students, Wang, Algozinne, Ma, and Porfeli (2011) report higher values and Marr, Algozzine, Nicholson, and Dugan (2011) report lower values. Finally, Kamata, Nese, Patarapicha-yatham, and Lai (2013) report higher values for 4th grade students.

The rate of change (e.g., one word per week) is very similar to many researchers who only reported slopes: a slope of 1.4 words per week (with 7 day week) by Tindal, Flick, and Cole (1992); slopes of 1.0 that decrease over grades are reported by Fuchs, Fuchs, Hamlett, Walz, and Germann (1993) and Deno, Fuchs, Marston, and Shin (2001), Hintz, Daley, and Shapiro (1998) (though their slopes were computed using a 5 day week so were inflated), and Silberglitt and Hintze (2007). Both Puranik, Petscher, Al Otaiba, Catts, and Lonigan (2008) and Logan and Petscher (2010) report similar slopes based on monthly growth. Ardoin and Christ (2008; 2009), Al Otaiba, Petscher, Pappamihiel, Williams, Dyrlund, and Connor (2009), and Christ, Silberglitt, Yeo, and Cormier, (2010) all reported similar slopes reflecting one word per week gain. Slightly higher slopes are reported by Jenkins, Graff, and Miglioretti (2009) as well as Jenkins and Terjeson (2011). The gain reported by Wanzek et al. (2010) was 11 wcpm from one time period to another over three years.

At the same time, some anomalies also exist and it is uncertain why. For example, the first three studies that reported fall, winter, and spring fluency rates showed considerable variation, particularly in Grades 1 and 2 (Marston, Lowry, Deno, & Mirkin, 1981; Tindal,

Germann, & Deno, 1983; Tindal, Germann, Marston, & Deno, 1983). The research by Simmons, Fuchs, Fuchs, Mathes, and Hodge (1995) and Fuchs, Fuchs, Mathes, and Simmons (1997) show unusually high values (primarily from averaging across students from different grade levels). Dunn and Eckert (2002) report daily growth over an 8-week period. Speece and Richey (2005) use 20 weeks to report their values; Jenkins, Zumeta, and Dupree (2005) also report slope over varying times while Jenkins and Terjeson, K. (2011) report changes at weeks 2, 4, and 8.

Finally, in the middle school grades, the most recent results from both Nese et al. (2013) and Tolar et al. (2014) reflect slightly higher values for initial performance (at least for typical students) to those reported earlier (Christ et al., 2010; Hasbrouck & Tindal, 1992, 2006; Yeo et al., 2011; Yeo et al., 2012). In Grade 6, students are reading approximately 150 *wcpm* at the beginning of the year and 175-180 at the end of the year (reflecting about one word growth per week). In Grades 7 and 8, the initial values are quite similar to the earlier studies but the eventual progress is slight (with little growth exhibited). Struggling students' performance is considerably lower though the growth is similar (at one-half to one word per week).

Summary

Research on oral reading fluency has been conducted over 30 years with considerable variation in methodology to address a number of questions about its technical adequacy: (a) sensitivity to different populations of students and generalizability of results, (b) identification of important variables that may influence performance and progress (measurement and student characteristics as well as treatments), (c) type of passages (both informally selected as well as formal standardized instruments), (d) frequency of measurement, and finally (e) the manner in which growth is documented. In the end, students' oral reading fluency within each grade level

and across successive grades has shown remarkable durability and stability in both performance and progress. The most significant differences appear in specific populations rather than in passages or administration conditions. Likewise, differences in the intercept between grades may be more significant than in the change over time. However, the values reported over successive decades are more similar than they are different, particularly if values are interpreted within confidence intervals.

Conclusions and Implications

After 30 years of research on oral (passage) reading fluency, the field of both general and special education now has an extensive database upon which to establish grade level expectations. Generally, we expect younger students to grow more within the year than older students; the possibility exists also that growth is not linear though the practical consequences may be minimal (e.g., some of these findings reflect statistical conclusions based on large sample sizes). Instrumentation has become more standardized and, in that sense, perhaps a bit more comparable. The findings reflect replications of performance levels that have occurred across different researchers, adding to the credibility of the findings. Though some variation exists across different studies, it may well be a function of the populations and the measures being studied (as well as time intervals). Even then, this variation is somewhat limited, particularly when considering reading rates as intervals for various grades.

Probably the most unfortunate (unintended) consequence is the degree to which the emphasis has been on establishing normative benchmark levels of performance and progress. Although this measurement system was designed for use in evaluating instructional programs with frequent progress measurement, the dominant trend has been its use to establish fall, winter, and spring benchmarks rather than use it in progress monitoring of instruction. Certainly some of

the studies listed in the table reflect evaluation of curriculum or instructional strategies but predominantly researchers are focusing on the scaling properties of the measures (standard error of measurement, reliability, criterion-related evidence, differences in materials and generalizability across populations, etc.). Perhaps this trend is to be expected in the natural cycle of scientific debate. That is, measures need to reflect technical adequacy prior to using them as a dependent variable for establishing cause-effect relations with independent variables.

To advance the field, however, it may be important to conduct more individually referenced studies in which changes are made to curriculum and instructional programs in synch with appropriate measurement of progress. This measurement also needs to be individually tailored to the student and not be restricted to the grade-level benchmark measures. In the ideal world of classroom use, then the benchmark (seasonal) measures would be used to identify the general goals for the student. Materials would then be selected that reflect these individual goals and not be restricted to simply grade level oral reading fluency. For example, progress measures may be selected from out of grade level passages or from early literacy measures or vocabulary measures. Once selected, alternate forms (from that grade level or suite of materials) would be maintained (administered consistently) over time to provide the trajectory of learning and the basis for evaluating the effects of instruction. At the same time, this instruction would be explicitly described with sufficient detail to systematically vary in determining not only what and how to teach but more importantly to understand why and with what effects. This model of hypothesis-driven instructional programs would in effect instantiate the scientific method in establishing cause effect relations using single case research designs (multiple baseline, multi-element treatment designs, etc.) and provide extensive replications across students.

References

Al Otaiba, S., Connor, C., Lane, H., Kosanovich, M. L., Schatschneider, C., Dyrlund, A. K., . . . Wright, T. L. (2008). *Reading First* kindergarten classroom instruction and students' growth in phonological awareness and letter naming-decoding fluency. *Journal of School Psychology, 46*, 281-314.

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

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

Ardoin, S. P., & Christ, T. (2009). Curriculum-based measurement of oral reading: Standard errors associated with progress montioring outcomes from DIBELS, AIMSweb, and an experimental passage. *School Psychology Review, 38*(2), 266-283.

Ardoin, S. P., Christ, T. J., & Morena, L. S. (2013). A systematic review and summarization of the recommendations and reseach surrounding Curriculum-Based Measurement f oral reading fluency (CBM-R) decision rule. *Journal of School Psychology, 51*, 1-18.

Baker, D. L., Park, Y., & Baker, S. K. (2012). The reading performance of English learners in grades 1-3: the role of initial status and growth on reading fluency in Spanish and English. *Reading and Writing, 25*, 251-281. doi: 10.1007/s11145-010-9261-z

Baker, S. K., & Good, R. (1995). Curriculum-based measurement of English reading with bilingual Hispanic students: A validation. *School Psychology Review, 24*(4), 561.

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'euni, E. J., & Thomas-Beck, C. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high poverty schools. *School Psychology Review, 37*(1), 18-37.

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

Christ, T., Silberglitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading (CBM-R): An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review, 39*(3), 447-462.

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

Cummings, K. D., Park, Y., & Bauer-Shaper, H. A. (2013). Form effects on DIBELS Next oral reading fluency progress-monitoring passages. *Assessment For Effective Intervention, 38*(2), 91-104. doi: 10.1177/1534508412447010

Deno, S., Mirkin, P., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36-43.

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*(4), 507-524.

Deno, S. L., & Mirkin, P. K. (1977). *Data based program modification: A manual*. Minneapolis, MN: Leadership Training Institute for Special Education.

Dunn, E. K., & Eckert, T. L. (2002). Curriculum-based measurement in reading: A comparison of simialr versus challenging material. *School Psychology Quarterly, 17*(1), 24-46.

Francis, D. J., Barth, A., Cirino, P. T., Reed, D., & Fletcher, J. M. (2008). The Texas middle school fluency assessment. Austin, TX: Texas Educational Agency.

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

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1982). Direct and frequent measuremet and evaluation: Effects on instructio and estimates of student progress. Minneapolis, MN: Institute for Research on Learning Disabilities-University of Minnesota.

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

Fuchs, L. S., Fuchs, D., Hamlett, C., 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., Tindal, G., & Deno, S. L. (1981). Effects of varying item domain and sample duration on technical charcteristics of daily measures in reading *Research Report*. Minneapolis, MN: Institute for Research on Learning Disabilities (IRLD)-University of Minnesota

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

Hasbrouck, J., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children, 24*(3), 41-44.

Hasbrouck, J., & Tindal, G. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*(7), 636-644. doi: 10.1598/RT.59.7.3

Hintze, J. M., Daly, E. J., III, & Shapiro, E. S. (1998). An investigation of the effects of passage difficulty level on outcomes of oral reading fluency progress monitoring. *School Psychology Review, 27*(3), 433-445.

Hintze, J. M., & Shapiro, E. S. (1997). Curriculum-based measurement and literature-based reading: Is curriculum-based measurement meeting the needs of chaning reading curricula? *Journal of School Psychology, 35*(4), 351-374.

Hintze, J. M., Shapiro, E. S., & Lutz, J. G. (1994). The effects of curriculum on the sensitivity of curriculum-based measurement in reading. *The Journal of Special Education, 28*, 188-202. doi: 10.1177/002246699402800205

Hosp, J. L., Hosp, M. K., Howell, K. W., & Allison, R. (2014). *The ABCs of curriculum-based evaluation: A practical guide to effective decision-making*. New York: The Guilford Press.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement (practical intervention in the schools)*. New York: The Guilford Press.

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

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

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

Kamata, A., Nese, J. F. T., Patarapichayatham, C., & Lai, C. F. (2013). Modeling nonlinear growth with three data points: Illustration with benchmarking data. *Assessment For Effective Intervention, 38*(2), 105-117. doi: 10.1177/1534508412457872

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

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

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293-323. doi: DOI: 10.1016/0010-0285(74)90015-2

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

MacMillan, P. (2000). Simultaneous measurement of reading growth, gender, and relative-age effects: many-faceted Rasch applied to CBM reading scores. *Journal of Applied Measurement, 1*(4), 393-408.

Marr, M. B., Algozzine, B., Nicholson, K., & Dugan, K. K. (2011). Building oral reading fluency with peer coaching. *Remedial and Special Education, 32*(3), 256–264. doi: 10.1177/0741932510362202

Marston, D., Lowry, L., Deno, S. L., & Mirkin, P. K. (1981). An analysis of learning trends in simple meaures of reading, spelling, and written expression: A longitudinal study. Minneapolis, MN: Institute for Research on Learning Disabilities-University of Minnesota.

Marston, D., & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. *Exceptional Children, 52*(3), 266-276.

National Assessment of Educational Progress. (2002). Oral Reading Study. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

National Reading Panel. (2000). Teaching children to read. Washington DC: National Institutes of Health.

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

Nese, J. F. T., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., & Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across grades 1-8. *Journal of School Psychology, 52*, 625-642.

Puranik, C. S., Petscher, Y., Al Otaiba, S., Catts, H. W., & Lonigan, C. J. (2008). Development of oral reading fluency in children with speech or language impairments : A growth curve analysis. *Journal of Learning Disabilities, 41*, 545-560.

Ramírez, R. D., & Shapiro, E. (2007). Cross-language relationship between Spanish and English orading fluency among Spanish-speaking English language learners in bilingual education classrooms. *Psychology in the Schools, 44*(8), 795-806.

Riley-Heller, N., Kelly-Vance, L., & Shriver, M. (2005). Curriculum-based measurement: Generic vs. curriculum-dependent probes. *Journal of Applied School Psychology, 21*(1), 141-162. doi: DOI:10.1300/J370v21n01_07

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

Shinn, M. (1989). *Curriculum-based measurement: Assessing special children*. New York: The Guilford Press.

Shinn, M. (1998). *Advanced applications of curriculum-based measurement*. New York: The Guilford Press.

Silberglitt, B., & Hintze, J. (2005). Formative assessment using CBM-R cut scores to track progress toward success on state-mandated achievement tests: A comparison of methods. *Journal of Psychoeducational Assessment, 23*, 304-325. doi: 10.1177/073428290502300402

Silberglitt, B., & Hintze, J. (2007). How much growth can we expect? A conditional analysis of R-CBM growth rates by level of performance. *Exceptional Children, 74*(1), 71-84.

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

Speece, D. L., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities, 38*(5), 387-399. doi: 10.1177/00222194050380050201

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30*(3), 407-419.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education, 2013*, 1-29. doi: 10.1155/2013/958530

Tindal, G., Flick, D., & Cole, C. (1992). The effect of curriculum on inferences of reading performance and improvement. *Diagnostique, 18*, 69-84.

Tindal, G., Germann, G., & Deno, S. L. (1983). Descriptive research on the Pine County norms: A compilation of findings. Minneapolis, MN: Institute for Research on Learning Disabilities (IRLD)-University of Minnesota

Tindal, G., Germann, G., Marston, D., & Deno, S. L. (1983). The effectiveness of special education: A direct measurement approach. Minneapolis, MN: Institute for Research on Learning Disabilities (IRLD)-University of Minnesota

Tindal, G., & Marston, D. (1990). *Classroom-based assessment: Evaluating instructional outcomes*. Columbus, OH: Merrill Publishing Company.

Tolar, T. D., Barth, A. E., Fletcher, J., Francis, D., & Vaughn, S. (2014). Predicting reading outcomes with progress monitoring slopes among middle grade students. *Learning and Individual Differences, 30*, 46-57.

Tucker, J. (1985). Curriculum-based assessment: An introduction. *Exceptional Children, 52*(3), 199-204.

Wang, C., Algozzine, B., Ma, W., & Porfeli, E. (2011). Oral reading rates of second-grade students. *Journal of Educational Psychology, 103*(2), 442–454.

Wanzek, J., Roberts, G., Linan-Thompson, S., Vaughn, S., L., W. A., & Murray, C. S. (2010). The relationship of oral reading fluency and high stakes measures of reading comprehension. *Assessment For Effective Intervention, 35*, 67-77.

Wayman, M. M., Wallace, T., Wiley, H. I., Tich'a, R., & Espin, C. A. (2007). Literature Synthesis on Curriculum-Based Measurement in Reading. *The Journal of Special Education, 41*(2), 85-120. doi: 10.1177/00224669070410020401

Wood, D. E. (2006). Modeling the relationship between oral reading fluency and performance on a statewide reading test. *Educational Assessment, 11*(2), 85-104.

Yeo, S., Fearrington, J., & Christ, T. (2011). An investigation of gender, income, and special education status bias on curriculum-based measurement slope in reading. . *School Psychology Quarterly, 26*(2), 119-130. doi: DOI: 10.1037/a0023021

Yeo, S., Fearrington, J. Y., & Christ, T. (2012). Relation between CBM-R and CBM-RM slopes: An application of latent growth modeling. *Assessment For Effective Intervention, 37*(3), 147-158. doi: 10.1177/1534508411420129

Table 1

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

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM			
Marston, Lowry, Deno, & Mirkin (1981)	Grades (n): 1 (13), 2 (9), 3 (10), 4 (7), 5 (7), 6 (9)	Students from a small Midwest city	Third grade basal reading series: Allyn-Bacon, Ginn 720, Houghton Mifflin	Three administrations: Fall, winter, and spring	Plot of raw scores and percentage of change; significance tests of differences	Grades	Fall	Winter	Spring
						1	18.1	31.1	45.7
						2	73.2	101.1	127.8
						3	108.3	123.6	136.2
						4	125.4	131.7	155.3
						5	125.7	147.3	161.1
						6	142.9	176.7	182.8
Tindal, Germann, Marston, & Deno (1983)	Grades (n): 1 (5), 2 (13), 3 (17), 4 (22), 5 (18), 6 (21)	Students from six districts referred, assessed, and eligible for special education	Curricula in use in the school district	Three administrations: Fall, winter, and spring	Change in raw score and in discrepancy from general education	Grades	Fall	Winter	Spring
						1	6.8	14.3	9.6
						2	5.5	16.2	23.6
						3	20.5	36.6	41.3
						4	31.0	50.3	52.9
						5	59.1	72.9	79.1
						6	59.5	65.9	66.8
Tindal, Germann, & Deno (1983)	Grades (n): 1 (276) 2 (284) 3 (302) 4 (294) 5 (315) 6 (328)	Students randomly sampled from six districts in Pine County	Two passages sampled from basal reading curriculum	Three administrations: Fall, winter, and spring	Change in raw scores	Grades	Fall	Winter	Spring
						1	5	63	75
						2	35	83	93
						3	67	89	108
						4	98	111	128
						5	121	120	138
						6	123	126	134
Fuchs, Deno, & Mirkin (1984)	Grades 3-5: 64 students in (DBPM) 77 students in no DBPM	All students were 'handicapped'	3rd grade passage reading test from Ginn 720	Pre-post (unknown: sometime between Nov. and May)	Pre-post difference @ 28 weeks	Condition*		Pre	Post
						Experimental*		41.6	70.2
						Contrast		51.5	51.3
						DBPM versus none			

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Marston & Magnusson, 1985)	1200 – 1500+ students in each of grade 1-6 from 35 schools in Minneapolis	Separated into regular and special education	Passages from Ginn 720	Tested by 65 trained teachers in three administrations: Fall, winter, and spring	Significant differences were found at each time period (F, W, S) between general (G) and special (S) education	Gr FG - FS WG - WS SG- SS 1 14.7 – 3.7 36.7 – 8.9 68.5 – 22.5 2 53.0 – 11.5 83.0 – 21.0 98.4 – 31.6 3 81.3 – 24.1 100.6 – 35.6 119.9 – 47.0 4 91.5 – 36.7 108.3 – 44.6 114.9 – 59.0 5 103.3 – 47.5 116.6 – 51.5 124.1- 64.4 6 129.8 – 58.8 138.0 – 63.2 143.4- 72.5 *F=Fall, W=Winter, S=Spring* *G=General Education, S=Special Education*
Hasbrouck & Tindal (1992)	Grades 3-5	Unspecified	Varied by site (school districts) used with data collection	Three administrations: Fall, winter, spring	Linear change across trimesters	Grades* Fall Winter Spring 2 53 78 94 3 79 93 114 4 99 112 118 5 105 118 128 **Reflects 50th percentile rank*
Tindal, Flick, & Cole (1992)	Grade 2 (6), Grade 3 (3), Grade 4 (2), and Grade 5 (1)	10 with learning disabilities, 1 with speech impairment, and 1 with other health impairment	Reading Mastery II and III, Distar 1, Merrill Linguistic Reader, and Scribner	Average of 1.4 to 2.6 measures per week across three phases of instruction	Pre-post difference on average equal to .15 to .61 across three phases of instruction	Baseline median of first three measures was 15 and median of final three measures was 59; the pre-post difference of 44 wcpm over 7 months equals 1.4 words per per week (with 7 day week)
Fuchs, Fuchs, Hamlett, Walz, & Germann (1993)	Grades 1-6: 3,057 over 2 years but slopes on only 374 in reading	Handicapped and non-handicapped students	'Generic reading passages were used' (p.31)	Unknown with 7+ measures for quadratic calculation	Ordinary Least Squares (OLS) for students with at least 7 measures throughout the year and multiplying change by 7 (to represent calendar weeks)	Grade: Count and weekly growth Grade 1 (n=33) = 1.2 Grade 2 (n=76) = 1.0 Grade 3 (n=66) = 0.8 Grade 4 (n=54) = 0.6 Grade 5 (n=77) = 0.5 Grade 6 (n=68) = 0.3
Hintze, Shapiro, & Lutz (1994)	48 students in grade 3 from two school districts (rural and urban)	Eliminated ELL, chapter 1, and special education students	Scott Foresman (literature-based) & Houghton Mifflin (traditional basal)	Twice per week for 9 weeks with average of 17 administrations	Ordinary Least Squares (OLS) regression with daily slope converted to weekly slope	Intercepts Lit. Inst. Trad Inst. Lit. Msr. 126.4 113.7 Trad. Msr 112.7 116.1 Weekly Growth Lit. Inst. Trad. Inst. Lit. Msr. -1.04 .66 Trad. Msr. -.34 1.72

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Simmons, Fuchs, Fuchs, Mathes, & Hodge (1995)	Grades 2-5: 21 Exp Tch 18 Exp Tch and Tutor 29 Control	Learning Disabled or Low Performer (< 25th PR)	*Ave. of two 3 min. ORFs: Comprehensive Reading Assessment Battery (folktales)*	Pre-post and Follow-up (FU)	Pre-post difference @ 16 weeks and FU @ 20 weeks	Condition Pre Post Follow ExpTch 181.9 223.1 212.7 ExpTchTutor 202.0 271.9 252.3 Control 159.5 196.9 203.7
Authors (date)	**Grades**	**Students**	**Measures**	**N-Measures**	**Slope Calculation**	**Averages and Growth – WCPM**
Baker & Good (1995)	Grade 2: 76 students from two schools	Core sample, English only and extended bilingual	20 passages from students' curriculum	Initial performance on day 1 and slope over time (unknown period)	Ordinary Least-Squares (OLS) regression line fit to the CBM data	Group Initial Slope English only 86.7 1.3 Bilingual 68.5 .1
Fuchs, Fuchs, Mathes, & Simmons (1997)	Average of 3rd grade for 66 students	Learning Disabled, Low Performers (< 25th PR), Average	*Comprehensive Reading Assessment Battery (400 word folktales)*	15 weeks of Peer Assisted Learning (PALS) or control	Pre-post difference using a one-between (treatment) and one-within (students) ANOVA	LD Pre Post Gain PALS 203.2 254.2 51.1 No PALS 202.2 230.9 26.7 LP Pre Post Gain PALS 243.5 290.8 56.3 No PALS 188.0 228.3 40.4 Average Pre Post Gain PALS 291.6 351.1 59.5 No PALS 311.0 348.4 37.4
MacMillan (2000)	1,691 students in grades 2-7 from 53 elementary schools (270-300 per grade)	No description provided	Unspecified (Guidebook for CBM)	Six reading probes at each grade and three norming periods in Oct., Nov., & Dec. with one passage per administration	Many faceted Rasch model using FACETS	Grade Ave. Fall Winter Spring Grade 2 71.4 103.4 125.2 Grade 3 87.3 101.4 111.3 Grade 4 91.2 101.1 107.7 Grade 5 93.6 100.2 106.2 Grade 6 95.2 100.2 104.5 Grade 7 96.4 100.0 103.6

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Hintz, Daley, & Shapiro (1998)	A total of 80 students enrolled in 12 first-, second-, third-, and fourth-grade classrooms	32% of the pupils receiving free or reduced lunch; ELL students eliminated; 43 males and 37 females of whom 80% were Caucasian; 88% received reading instruction in general education	20 reading probes per grade level for both grade and goal level material for grades 1-4 in the reading series being taught; goal level material represented material in which students are expected to be proficient in 1 year.	Progress-monitoring sessions were conducted twice a week during a 10-week period. As such, each student had a maximum of 20 sessions.	Ordinary least squares (OLS) regression for each type of material with slope converted to weekly improvement (x5)	Slope of growth over 10 weeks Grade / Grade / Goal Gr 1 / 3.3 / 2.0 Gr 2 / 0.7 / 0.3 Gr 3 / 0.2 / 0.1 Gr 4 / 1.6 / 1.9 All / 1.4 / 1.0
Stage & Jacobsen (2001)	173 fourth grade students from one elementary school	54% male and 46% female with 11 in special education	Silver Burdett and Ginn passages from reading curriculum	Three administrations: September, January, and May	Hierarchical linear analysis modeling (HLM) to calculate slope with ORF and with WA state test	Word Correct/Minute Fall / Winter / Spring 107.3 / 122.1 / 136.9
	Grades 1 – 6 with 2,999 students (urban north, urban south, rural mid-west, and mid-size west)	General and special education	'Grade-appropriate' as determined by LEAs	Weekly and seasonally (fall, winter, spring)	Ordinary Least Squares (OLS) regression	Intercept / General / Special Grade 1 / 0.0 / 0.0 Grade 2 / 33.7 / 11.8 Grade 3 / 69.9 / 21.8 Grade 4 / 86.0 / 28.4 Grade 5 / 104.2 / 45.1 Grade 6 / 111.2 / 44.6 Weekly Growth / General / Special Grade 1 / 1.8 / .8 Grade 2 / 1.7 / .6 Grade 3 / 1.2 / .6 Grade 4 / 1.0 / .6 Grade 5 / .6 / .6 Grade 6 / .7 / .6

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Dunn & Eckert (2002)	20 students (male, *n* = 11, female, *n* = 9) from grades 2 and 3	Grade 2 instructional level with no additional instruction in reading	Passages from current instruction (similar) versus passages from materials expected to achieve in a year (challenging)	32 passages read over 8 weeks	Ordinary least square (OLS) estimate of the slope line based on each student's WRCM over the 8-week period	Daily Growth: Similar / Challenge Group 1: .2 / .2 Group 2: 1.0 / .2 Total: .1 / .2 *Weekly Growth: .7 / .9
Speece & Richey (2005)	Grade 1: 140 at risk 136 not at risk	Student gender and race-ethnicity	School district reading curriculum	Administration over 20 weeks (weekly for 6 weeks and then monthly from January to May)	Growth curve analysis to identify predictors of oral reading fluency	Group: Intercept / Slope At Risk: 6.4 / .8 Not At Risk: 11.1 / 1.6
Silberglitt and Hintz (2005)	2,100 students from a MN educational cooperative (1400 – 2100 students in grade 1-3	Student gender and race-ethnicity summarized for group	CBM-R passages (no other reported information)	Three administrations: Fall, winter, and spring	Student performance was highly predictive of the state test in identifying students at risk of reading failure	Grade: Fall / Winter / Spring Grade 1: 29.0 / 56.4 Grade 2: 48.3 / 73.4 / 91.5 Grade 3: 73.4 / 95.0 / 109.9
Plascenia-Peinado, Deno, & Johnson (2005)	Three low, middle, and high scoring students from 134	First grade students with unknown demographics	Alternate forms of grade 1 passages	Six weeks of weekly administration	Differences in oral reading growth with high performers showing lowest slope; correction	Grp-Slope: Wk1 / Wk3 / Wk6 Low (2.1): 13.4 / 19.4 / 27.4 Middle (2.3): 34.8 / 49.8 / 52.9 High (1.4): 73.2 / 82.6 / 82.3

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Jenkins, Zumeta, & Dupree (2005)	Grade (count) Grade 2 (1) Grade 3 (4) Grade 4 (9) Grade 5 (6) Grade 6 (9) Grade 7 (3) Grade 10 (1)	33 students attending a private school for students with learning disabilities (mostly white males)	Ten readability controlled reading passages from Edcheckup	Three administrations: Sept. through Jan. (10 weeks apart from test 1-2 and 5 weeks apart from test 2-3); students read from 4 passages.	Gains from time to time using raw scores	10 weeks same: Same / Diff Pair 1: 16.6 / 10.8 Pair 2: 11.4 / 9.8 Pair 3: 15.5 / 12.1 Pair 4: 8.7 / 14.1 Overall: 13.0 / 11.7 5 weeks same: Same / Diff Pair 1: 6.6 / 5.5 Pair 2: 7.7 / 0.6 Overall: 7.2 / 3.0
Ramirez & Shapiro (2006)	Grades K – 5: 68 students in bilingual programs	Transitional bilingual programs with no special education students	AIMSweb CBM-R, and Spanish text from trade books	Three administrations: Fall, winter, and spring	Pre (fall) to Winter and Spring (post) test performance	Grade Ave. / Fall / Winter / Spring Grade 1: 6.8 / 14.8 / 26.1 Grade 2: 40.1 / 50.7 / 58.2 Grade 3: 61.5 / 69.2 / 77.6 Grade 4: 61.5 / 67.7 / 75.9 Grade 5: 89.0 / 103.8 / 116.6
Wood (2006)	Grades 3-5: 82, 101, and 98 students, respectively	Students tracked by five classrooms (level 2); race and special education noted for group only	**DIBELS** reading passages at grade level	Winter benchmark administered in December	Hierarchical linear model (HLM) with growth over grades	Grade / Mean / Classroom Range 3: 103: 92.6 to 112.6 4: 116: 108.4 to 119.5 5: 135: 120.0 to 154.5 **A significant increase in fluency each year was 16.4 words per minute from Grade 3 to 4 and from Grade 4 to 5*
Hasbrouck & Tindal (2006)	Grades 1-8: Range from 11,00-20,000 students (grades 1-6) and 5,000 to 6,000 students (grades 7-8)	Unspecified	Varied by site (school districts) used with data collection	Three administrations: Fall, winter, and spring	Linear change across trimesters	Grades / Fall* / Winter* / Spring* 1: - / 23 / 53 2: 51 / 72 / 89 3: 71 / 92 / 107 4: 94 / 112 / 123 5: 110 / 127 / 139 6: 127 / 140 / 150 7: 128 / 136 / 150 8: 133 / 146 / 151 **Reflects 50th percentile rank*

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Silberglitt & Hintze (2007)	Grades 2-6: 7,544 students from five districts in MN	Divided students into deciles (relative to 50th to 59th PR as the reference)	Two sets of probes: Silver, Burdett, Ginn and AIMSweb	Three administrations: Fall, winter, and spring (18 weeks apart)	Hierarchical linear model (HLM): level 1 is time within student and level 2 is student decile in fall; weekly growth assumed to be the metric though no descriptive statistics are presented	Grade 2: 50-59 PR 1.3, Decile Range -.5 to .10 Grade 3: 1.1, -.4 to .10 Grade 4: .9, -.1 to .04 Grade 5: .9, -.1 to .03 Grade 6: .8, -.1 to .00 **Growth is significantly greater in earlier than in later grades.*
Christ & Silberglitt (2007)	Grades 1-5: 4,196 to 5,036 for a total of 8,200	Gender and race reported for group only	AIMSweb CBM-R passages	Three administrations: Fall, winter, and spring from 1996 to 2004	Raw score change	Grade 1: Winter 32, Spring 60 Grade 2: Fall 50, Winter 77, Spring 95 Grade 3: Fall 74, Winter 95, Spring 108 Grade 4: Fall 93, Winter 113, Spring 125 Grade 5: Fall 112, Winter 130, Spring 139
Schatschneider, Wagner, & Crawford (2008)	Grade 1: 23,438 students	Students attending Reading First schools and primarily White, Black, and Hispanic.	**DIBELS** oral reading fluency	ORF was administered four times during the months of September, December, February, and April.	Compare predictive validity of estimates of (a) student growth in oral reading fluency, (b) student status or level of oral reading fluency, and (c) combined measures of growth and status, for prediction of concurrent and future reading skills	Ave. Sep. 13.7, Dec. 20.3, Feb. 34.3, Apr. 49.6

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Puranik, Petscher, Al Otaiba, Catts, & Lonigan (2008)	Grades 1-3: 1,991 students with 3 years of data and were in 1st grade in 2003-2004	1,388 Speech Impaired and 603 Language Impaired classified as (a) Resolved, (b) Persistent, I became LD	**DIBELS** (Kaminski & Good, 1996)	ORF was administered four times per year (i.e., September, December, February, April)	Hierarchical piecewise growth curve model (PGCM) across first, second, and third grades	Monthly Growth Population Gr 1 Gr 2 Gr 3 SI-PE .49 -2.1 5.1 LI-PE -.11 -1.9 4.6 SI-RE .89 -1.9 5.1 LI-RE 1.02 -3.7 4.0 SI-LD .42 -2.4 4.5 LI-LD .11 -1.6 3.7 Norm .49 -1.8 5.5
Baker, Smolkowski, Katz, Fien, Seeley, Kame'enui, & Beck (2008)	14, 495 students from various cohorts: 4,973 grade 1 4,826 grade 2 4,696 grade 3	34 schools eligible for Reading First (based on poverty and reading performance)	**DIBELS** (Kaminski & Good, 2002), SAT-10, and Oregon Statewide Reading Assessment	Three administrations: Fall (Sept.), winter (Jan.), and spring (April)	Growth curve analyses to address ORF trajectories (intercepts and slopes) in predicting performance on SAT-10 or OSRA (end of Year 2)	Cohorts Fall Wint Spring Gr 1-Yr 2 24.1 45.7 Gr 1-Yr 1 20.5 41.3 Gr 2-Yr 2 37.2 63.1 80.2 Gr 2-Yr 1 32.8 58.0 74.9 Gr 3-Yr 2 62.5 79.6 97.5 Gr 3-Yr 1 58.4 76.5 94.1
Ardoin & Christ (2008)	Grades K – 5: 540 students	Gender and race reported only for group	**DIBELS** 2nd grade passages	Three administrations: Fall (Sept.), winter (Jan.), and spring (April)	A 2 (semester) by 3 (growth estimate across probe-set combinations) repeated multivariate analysis (MANOVA)	Probe* Fall Winter Spring Prog. Mon. 87 111 119 Bench1 90 115 126 Bench2 78 97 108 Bench3 74 96 110 **Only includes those given 3 times* Growth* F-W W-S F-W-S Bench1 1.3 .6 .9 Bench2 1.2 .9 1.0 Bench3 1.6 .4 1.0 *F=Fall, W=Winter, S=Spring* **Only includes benchmarks*

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Riley-Heller, Kelly-Vance, & Shriver (2008)	13 students from grade 2	7 males and 6 females eligible for free-reduced price lunch, 40% Latino and 60% white; no ELL	Phonics-based curriculum (Foresman) being taught (curriculum dependent program (CDP) and Test of Reading Fluency (TORF)	Three passages administered twice each week for 5 weeks	Ordinary least square (OLS) multiple regression	Day-Week / CDP / TORF / Ave. T-wk1 38.4 41.4 39.9 Th-wk1 40.0 42.7 41.4 T-wk2 41.8 44.3 43.1 Th-wk2 43.5 44.7 44.1 T-wk3 45.5 46.7 46.1 Th-wk3 46.2 49.4 47.8 T-wk4 48.2 54.2 51.2 Th-wk4 48.7 57.3 53.0 T-wk5 56.3 59.2 57.8 Th-wk5 53.7 58.8 56.3
Graney, Missall, Martinez, & Bergstrom (2009)	Grades 3 – 5: 442 in Yr1 456 in Yr2	Gender, race and subsidized lunch reported only for groups N.B. Half population had special education status	AIMSweb CBM-R passages from 2004-2006	Three administrations: Fall, winter, and spring with 19 weeks between fall-winter and 12 weeks between winter and spring	Raw score difference / number of weeks in interval (e.g., 19 and 12)	Year 1 / Fall / Winter / Spring Grade 3 GE 93 108 118 Grade 3 SE 55 60 73 Grade 4 111 120 133 Grade 5 121 129 141 Year 2 / Fall / Winter / Spring Grade 3 GE 79 101 116 Grade 3 SE 46 65 80 Grade 4 GE 113 130 143 Grade 4 SE 76 89 102 Grade 5 117 132 146 **Weekly growth = .41 to 1.1 (Fall) and .79 to 1.14 (Spring) in both years*
Jenkins, Graff, & Miglioretti (2009)	Grades 3 – 8: 41 students	All LD; gender, race, and ELL reported for group	Standard reading passages from Vanderbilt	A total of 29 measures in 1, 2, 3, 4 weeks and pre-post (Sept.-Nov.)	Ordinary least square (OLS) linear regression	True slope growth=1.1 One BL and every 1 week = 1.5 One BL and every 2 weeks = 1.9 One BL and every 3 weeks = 1.8 One BL and every 4 weeks = 1.8 One BL and post = 1.6

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Al Otaiba, Petscher, Pappamihiel, Williams, Dyrlund, & Connor (2009)	Grades 2- 3: 5,004 Latino students	(a) Proficient in English, (b) Not proficient and receiving English as a second language (ESL) services, and (c) Proficient enough to have exited from ESL	**DIBELS** (Good & Kaminski, 1996)	Four administrations: first 20–30 days of school (Sep); between the 65th and 75th days of school (Nov); between the 110th and 120th days of school (Feb); and between the 155th and 165th days of school (Apr).	Two level HLM with growth centered on the first testing time in third grade (September) and student characteristics (i.e., language group and special education subgroup) were entered at Level 2 designed to model both second- and third-grade growth trajectories	Grade 2 Weekly Growth: Proficient ESL ESL-exit GE 1.2 .8 1.2 LD 1.2 .8 1.1 SL 1.3 .9 1.1 Grade 3 Weekly Growth: Proficient ESL ESL-exit GE 1.2 1.1 1.3 LD 1.3 1.1 1.3 SL 1.2 1.1 1.2
Ardoin & Christ (2009)	Grades 2-3: 28 and 40 respectively	Race and free or reduced-price lunch reported for students in each of two schools	Three passage sets: (a) FAIP-R, (b) AIMSweb, and (c) **DIBELS**	12 weeks (with approximately 2 administrations per week) with time of year unknown	Ordinary least square (OLS) regression used to calculate an intercept and slope for each passage set	Intercept Slope* FAIP-R 84.8 1.4 AIMSweb 94.3 1.8 DIBELS 100.4 .5 **Slope times 7 to estimate weekly gain*
Crowe, Connor, & Petscher (2009)	Grade 1 (n=9,993), Grade 2 (n=9,869), Grade 3 (n=10,141)	Lower SES (eligible for free or reduced price lunch) and non lower SES students	**DIBELS** (Good & Kaminski, 2002).	Two administrations in September and April	Hierarchical Linear Modeling (HLM) to estimate mean growth trajectories for curriculum interacting with SES over the school year (7 months)	Sept. April Grade 1 18 50 • Higher SES 25 63 • Lower SES 16 46 Grade 2 53 89 • Higher SES 63 77 • Lower SES 50 62 Grade 3 73 101 • Higher SES 85 113 • Lower SES 69 97 **Curriculum differences reported that interacted with students SES.*

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Logan & Petscher (2010)	Grades 1-3: Grade 1 (58,844), grade 2 (56,768), and grade 3 (57,873)	Gender, race, and free-reduced price lunch reported for group only. Risk defined as minority status, English proficiency, and lunch status.	**DIBELS** (Good & Kaminski, 2002).	Four administrations in September, December, February, and April in 2005-2006 academic school year	Three-level growth curves were fit and latent profile analysis used to cluster schools into groups based on their percent of students meeting three risk categories (minority, ELL, and FRL students within each school) with six models compared	Intercept & Monthly Growth Low-Risk (intercept = 18) = 4.0 Average-Risk (intercept = 17) = 3.5 Poverty-Risk (intercept = 15) = 3.2 Language-Risk (intercept = 14) = 3.1
Christ, Silberglitt, Yeo, & Cormier, (2010)	Grades 2 – 6: 4,824 students	Gender and race reported for group only	AIMSweb CBM-R passages from 2001-2005	Three administrations: Fall, winter, and spring passages (0, 18, and 36 weeks)	Linear Mixed Model (LMM) for linear and piece wise growth	Grade GE* Fall Winter Spring 2 57 88 107 3 82 107 125 4 106 128 141 5 125 145 160 6 142 158 171 Grade SE* Fall Winter Spring 2 33 55 74 3 52 72 88 4 77 96 108 5 89 106 118 6 106 119 132 *Growth Fall-Wint Wint-Spring GE .9 to 1.7 .7 to 1.0 SE .7 to 1.2 .7 to 1.1
Kim, Petscher, Schatschneider, & Foorman (2010)	Grades 1-3: 13,154 enrolled in a Reading First school over 4 years	Gender, race, and free-reduced price lunch reported for group only	**DIBELS** (Good, Kaminski, Smith, Laimon, & Dill, 2001)	Three administrations: Fall, winter, and spring	Five multilevel growth models were used to estimate students' average level of performance at each grade (i.e., initial or end-of-year status), average rate of change for each predictor on their reading comprehension achievement; dominance analysis was used to compare predictors	Grade Fall Wint1 Wint2 Spr 1 22.5 35.1 45.5 55.5 2 59.2 72.8 85.8 96.9 3 86.1 101.1 121.3 **Per month growth: Gr 1 = 3.0, Gr 2 = 4.4, and Gr 3=5.6* *1st gr ORF growth most predictive*

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Wanzek, Roberts, Linan-Thompson, Vaughn, S., Woodruff, & Murray (2010)	461 students followed from grade 1 through grade 3	Gender, race, and disability reported for group only	DIBELS; (Good & Kaminski, 2002)	Three administrations: Fall, winter, and spring	Latent factors of linear fluency growth (using full information maximum likelihood) as predictors of SAT-10 & Texas Assessment of Knowledge & Skills	Grade Fall Winter Spring 1 30.0 52.2 2 46.0 77.4 88.9 3 73.9 88.3 102.2 ** Average of 11.33 words correct per minute per time point*
Yeo, Fearrington, & Christ (2011)	Grades 3 – 8: 1,738 in 2 elementary and 3 middle schools	Students coded for gender, race, free lunch, and special education status	AIMSweb CBM-R passages administered in 2006-2007	Three administrations: Fall, winter, and spring	Latent growth modeling (LGM) with three models: (a) no growth, (b) linear growth, and (c) conditional growth (free lunch, gender, special education). Sig. predictors found in grades 3 and 7	Averages Fall Winter Spring Grade 3 68.8 92.7 106.4 Grade 4 91.5 111.9 123.9 Grade 5 105.3 117.3 130.3 Grade 6 120.7 130.9 142.5 Grade 7 128.5 137.4 150.7 Grade 8 126.7 141.6 149.6 **Growth rate of CBM-R scores from fall to spring were homogenous among subgroups, except for grades 3 and 7.*
Jenkins, J., & Terjeson, K. (2011)	Grades 2 – 6: 31 students	All students with a disability at risk or with some risk	**DIBELS** passages	Measures every 2, 4, and 8 weeks	Ordinary least squares (OLS) slopes	Time Level Baseline 52.8 Week 2 55.5 Week 4 59.6 Week 6 65.2 Week 8 64.7 Slopes Every 2 weeks = 1.7 Every 4 & 8 weeks = 1.5 Every 3 & 8 weeks = 1.3 Every 5 & 8 weeks = 1.6
Wang, Algozinne, Ma, & Porfeli (2011)	Grade 2: 5,796 second-grade students in a large urban public school system in North Carolina	Gender, race, ELL status, and free reduced-price lunch reported for group only.	**DIBELS** passages (Good & Kaminski, 2002)	Three administrations: Fall, winter, and spring	Growth curve analysis was completed using Hierarchical Linear Model (HLM)	Gen Ave. Fall Winter Spring Male 61.1 87.9 103.3 Female 69.5 97.5 111.6 Sped Ave. Fall Winter Spring Male 47.2 68.6 84.8 Female 51.9 76.3 89.3

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Marr, Algozzine, Nicholson, & Dugan (2011)	Grade 2: 14 elementary schools	17 coached 17 control	**DIBELS** (Good & Kaminski, 2002)	Three administrations: Fall, winter, and spring	Ordinary Least Squares (OLS): One between-one within ANOVA Follow up t-tests	Averages: Control Fall 32.5, Winter 51.7, Spring 70.7; Treatment Fall 43.5, Winter 79.5, Spring 96.9 Growth: Control Fall-Wint 19, Wint-Spring 19; Treatment Fall-Wint 35, Wint-Spring 17
Keller-Margulis, Clemens, Im, Kwok, & Booth (2012)	Grades 3-5: Grade 3 (n=1,838), Grade 4 (n=2,151), Grade 5 (n=2,332)	Non-ELL, ELL, and ELL-Monitor	AIMSweb CBM-R passages (2012)	Three administrations: September, January (14 weeks later), and May (14 weeks later).	A piecewise model was fitted to the data that modeled growth	Averages: Grade 3 Fall 82.7, Winter 102.8, Spring 118.0; Grade 4 Fall 99.1, Winter 115.8, Spring 129.5; Grade 5 Fall 112.3, Winter 126.5, Spring 140.0 Growth: Gr 3: Non-ELL Fall-Wint. 20, Wint.-Spring 15; Gr 3: ELL 19, 17; Gr 3: ELL-M 20, 15; Gr 4: Non-ELL 17, 13; Gr 4: ELL 15, 12; Gr 4: ELL-M 17, 15; Gr 5: Non-ELL 15, 13; Gr 5: ELL 15, 13; Gr 5: ELL-M 14, 15
Nese, Biancarosa, Anderson, Lai, Alonzo, & Tindal (2012)	Grades 3 to 5: 2,465 students	Student gender, free and reduced-price lunch, ethnicity, special education status, and limited English proficiency	easyCBM passages	Three administrations: Fall, winter, and spring	Hierarchical Linear Model (HLM) for linear and discontinuous growth	Averages: 3* Fall 74.5, Winter 106.9, Spring 107.5; 4* Fall 102.3, Winter 122.9, Spring 130.9; 5 Fall 134.5, Winter 143.4, Spring 156.0 Growth: 3* Fall-Wint 32, Wint-Spring 2; 4* 24, 8; 5 10, 14 **Reflects curvilinear growth. N.B. Differences in intercept by student demographics and in slope for SWD.*

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Baker, Park, & Baker (2012)	Grades 1-3: 471 English language (EL) learners from 2006-2008	EL specifically and ethnicity and free-reduced price lunch for the group	**DIBELS** passages (Good & Kaminski, 2002)	Three administrations: Fall, winter, and spring	Hierarchical multivariate linear models with English as reference group to estimate the average initial status and growth rate of Spanish and English fluency within participants	Grade 2: Fall 28.2, Winter 55.4, Spring 74.1 Grade 3: Fall 53.9, Winter 77.1, Spring 99.7 ** Scores are ORF in English: ELs in second and third grades made more growth on oral reading fluency in English than in Spanish.*
Kamata, Nese, Patarapicha-yatham, & Lai (2013)	Grade 4 (n): Time 1 (2,166) Time 2 (2,216) Time 3 (2,194)	No description of students	Grade level easyCBM benchmark passages	Three administrations Fall, winter, and spring	Compared: (a) linear growth (b) piece-wise growth mixture model, (c) growth mixture modeling	Fall 110.0, Winter 132.4, Spring 141.3 • Change Fall to Winter = 22.4 • Change Winter to Spring = 8.9
Yeo, Fearrington, & Christ (2012)	Grades 3-8: 1,528 students from two elementary and three middle schools	Description of race-ethnicity, special education, ELL, and free reduced-price lunch for group only	AIMSweb passages	Three administrations Fall, winter, and spring	Documented and correlated growth on CBM-R (ORF) with CBM-Rm (maze) and state test	Grade 3: Fall 70.3, Winter 93.8, Spring 107.3 Grade 4: Fall 92.7, Winter 111.9, Spring 124.3 Grade 5: Fall 106.6, Winter 119.2, Spring 131.3 Grade 6: Fall 122.5, Winter 131.6, Spring 142.5 Grade 7: Fall 130.9, Winter 138.8, Spring 151.6 Grade 8: Fall 126.8, Winter 141.2, Spring 149.2
Nese, Biancarosa, Cummings, Kennedy, Alonzo, & Tindal (2013)	Grades 1-8: $n \approx 1{,}493$	Description of gender, race-ethnicity, ELL, and special education only	Grade level easyCBM benchmark passages	Grade 1 with 6 passages; Grades 2-8 with 8 passages	Uses latent growth modeling to determine quadratic (decelerating) growth model fit better in most grades rather than linear or cubic and that individual varying weeks, not average weeks or months, fit best.	O N D J F M A M J 1 19 - - 46 48 57 63 69 - 2 70 89 100 91 107 118 115 109 - 3 96 118 126 131 126 131 140 131 - 4 122 134 145 145 144 148 144 153 - 5 158 166 172 169 176 165 186 - 182 6 152 148 160 166 170 172 169 182 - 7 146 148 141 156 144 159 157 - 150 8 158 154 146 148 156 157 159 - 151 *O=Oct, N=Nov, D=Dec, J=Jan, F=Feb, M=Mar, A=April, M=May, J=June*

Authors (date)	Grades	Students	Measures	N-Measures	Slope Calculation	Averages and Growth – WCPM
Tolar, Barth, Fletcher, Francis, & Vaughn (2014)	1,343 middle school students (grades 6–8)	Description of gender, race-ethnicity, and free or reduced lunch. Organized into three groups: adequate, struggling no intervention, struggling with intervention	Oral Reading Fluency CBM-Passage Fluency (Francis, Barth, Cirino, Reed, & Fletcher, 2008). Familiar and novel progress monitoring passages	Five administrations, once every two months using mean of 3 administrations each time	Significant linear growth in ORF-PF for all groups. Significant variance in linear growth among all familiar groups. No significant variance in growth among the novel groups. ORF-PF slope was generally not a significant predictor of reading achievement (only for typical students).	T1 T2 T3 T4 T5 FT 145.4 157.3 165.4 171.4 174.7 FSnI 109.8 122.0 127.6 133.5 138.1 FSI 113.2 125.7 131.8 136.7 142.9 NT 147.1 155.7 158.4 161.7 162.5 NSnI 114.3 125.6 125.9 126.4 127.8 NSI 116.1 124.1 125.4 128.4 130.8 *T=Time Intervals (1-5)* *F=Familiar and N = novel* *SnI=Struggling with no intervention* *SI = Struggling with intervention*