Journal of Education for Students Placed at Risk (JESPAR)
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653666

Development of Oral Reading Fluency in Young Children at Risk for Failure

Chuang Wang [a]; Erik Porfeli [b]; Bob Algozzine [a]
[a] Department of Educational Leadership, University of North Carolina at Charlotte, [b] Department of Behavioral Sciences, Northeastern Ohio Universities College of Medicine,

Online Publication Date: 01 October 2008

To cite this Article Wang, Chuang, Porfeli, Erik and Algozzine, Bob(2008)'Development of Oral Reading Fluency in Young Children at Risk for Failure',Journal of Education for Students Placed at Risk (JESPAR),13:4,402 — 425
To link to this Article: DOI: 10.1080/10824660802427702
URL: http://dx.doi.org/10.1080/10824660802427702

Journal of Education for Students Placed at Risk, 13:402–425, 2008

ISSN: 1082-4669 print/1532-7671 online
DOI: 10.1080/10824660802427702

Routledge
Taylor & Francis Group

Development of Oral Reading Fluency in Young Children at Risk for Failure

Chuang Wang
Department of Educational Leadership
University of North Carolina at Charlotte

Erik Porfeli
Department of Behavioral Sciences
Northeastern Ohio Universities College of Medicine

Bob Algozzine
Department of Educational Leadership
University of North Carolina at Charlotte

The body of knowledge on critical components of reading instruction is broad and deep with phonemic awareness, phonics, fluency, vocabulary, and comprehension serving as building blocks for literacy. The importance of fluent reading (i.e., with speed, accuracy, and proper expression, as well as comprehension) in the formula for success is widely accepted. The purpose of this study was to evaluate oral reading rates in a sample of 2nd-grade children. Using growth curve analysis, we identified models for a combined sample of at-risk and not-at-risk children. Large and important differences in oral reading were evident, but the growth rates remained the same throughout the 2nd-grade; in fact, the gap between at-risk and not-at-risk students remained constant. We discuss the implications of our findings for future research and the improvement of practice.

The body of knowledge about how to teach reading is broad and grounded in solid support for systematic and explicit beginning literacy instruction

Request for reprints should be sent to: Chuang Wang, BRIC/EDLD/COED, University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223. E-mail: cwang15@uncc.edu

addressing the importance of the alphabetic principle and phonemic awareness, explicit phonics instruction, practice in building fluency, the development of vocabulary, and instruction in comprehension strategies and skills (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Kuhn et al., 2006; Mesmer, 2006; National Institute of Child Health and Human Development, 2000a, 2000b, 2000c; National Reading Panel, 2000; National Research Council, 1998). Developing an understanding of the alphabetic principle—that sounds can be mapped onto letters—is necessary as children connect and blend letters and sounds to read words. At the same time, developing phonemic awareness, the knowledge that words are made up of sounds, is critical to learning to read and spell successfully. Explicit phonics instruction provides children with a powerful strategy to decode written language and to read or access unfamiliar words confidently as they encounter them in text. Rapid, accurate, and expressive (i.e., fluent) word recognition builds vocabulary and serves as the basis for achieving comprehension, the ultimate goal of reading. Inadequate development of these basic early literacy skills represents a failure that is difficult to overcome.

Fluency has been at the top of the list of important topics for literacy researchers for some time (Cassidy & Cassidy, 2003, 2004a, 2004b, 2005, 2007; Cassidy, Garrett, & Barrera, 2006). It represents "the skill of the hour" (Marsell, 2007, p. 8) and its importance is widely accepted and supported (Abadiano & Turner, 2005; Behavioral Research and Teaching, 2004; Chard, Pikulski, & McDonagh, 2006; Chatel, 2005; Cunningham & Stanovich, 1997; Dudley & Mather, 2005; Ellery, 2004; Foorman & Moats, 2004; Good, Wallin, Simmons, Kame'enui, & Kaminski, 2002; Hasbrouck & Tindal, 1992; Johns & Berglund, 2005; Johnstone, Moen, Thurlow, Matchett, Hausmann, & Scullin, 2007; Kame'enui & Simmons, 2001; Kuhn, 2004; Kuhn & Stahl, 2000; Moats, 2001; Oakley, 2005; Pikulski & Chard, 2005; Prescott-Griffin & Witherell, 2004; Rasinski, 2000, 2003, 2005; Rasinski, Blachowicz, & Lems, 2006; Reutzel, 2006; Samuels, 1997, 2006; Samuels, Ediger, & Fautsch-Patridge, 2005; Shanahan, 2006; Stage & Jacobsen, 2001; Stage, Sheppard, Davidson, & Browning, 2001; Starch, 1915; Walker, Mokhtari, & Sargent, 2006; Wilson, 2005; Witte-Townsend & Whiting, 2005). Fuchs, Fuchs, and Maxwell (1988) reported high positive correlations between oral reading rates and comprehension. Jenkins, Fuchs, van den Broek, Espin, and Deno (2003) reported similar predictive relationships. Schilling, Carlisle, Scott, and Zeng (2007) found that oral reading fluency as measured by the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) significantly predicted end-of-grade achievement on the Reading Total subtest of the Iowa Test of Basic Skills.

Progress monitoring is the scientifically based practice of assessing academic performance on a regular basis (weekly or monthly) to determine whether children are profiting appropriately from the typical instructional

program and to build more effective programs for those who are not (Fuchs & Fuchs, 2006). Because oral reading rates are good predictors of performance on more generalized outcome measures such as high-stakes achievement tests, they have become popular markers in progress monitoring efforts. Teachers use progress monitoring to make educational decisions and improve instructional effectiveness. These assessments provide frequent data for use in making decisions about an individual student's instructional needs. For example, based on a student's progress, a teacher may decide to increase the amount of instruction, slow the pace of the instruction, or change the instructional method completely. The use of progress monitoring instruments in special education has been shown to improve student outcomes in academic areas and has been a widely accepted evidence-based practice for many years (e.g., Fuchs, 2004; Fuchs & Fuchs, 1996; Fuchs, Fuchs, & Hamlett, 1989; Madelaine & Wheldall, 2004; Safer & Fleischman, 2005).

Much is known and professed about the development of early literacy skills related to the alphabetic principle, phonemic awareness, and explicit phonics instruction as well as regarding the value of fluent reading for vocabulary development and comprehension (Daane, Campbell, Grigg, Goodman, & Oranje, 2005; Ehri, Nunes, Stahl, & Willows, 2001; Foorman & Torgesen, 2001; National Reading Panel, 2000; Speece & Ritchey, 2005). Oral reading fluency norms for students in second to fifth grade have been published (cf. Hasbrouck & Tindal, 1992) and updated from 2000–2004 to represent "current student performance" (Hasbrouck & Tindal, 2006, p. 637). The purpose of the *Fourth-Grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading* was "to examine aspects of oral reading performance—accuracy, rate, and fluency—that cannot be observed from results of the main National Assessment of Educational Progress (NAEP) reading assessment" (Daane et al., 2005, p. iii). In 2002, after sitting for the main NAEP reading assessment, 1,779 (1%) of the 140,000 fourth graders read aloud a 198-word excerpt from one of the passages they encountered 1 week earlier and their performance was compared across gender and racial/ethnic groups. In terms of accuracy, rate, and fluency, girls performed better than boys and Caucasian students outperformed their Hispanic and African American peers. In this research, we were interested in oral reading fluency as an early marker for achievement patterns in America's schools. We investigated similarities and differences in oral reading rates within and across diverse groups of children.

METHOD

The importance of fluent reading is widely accepted, but there are few data illustrating and comparing oral reading fluency across progress monitoring

benchmarks for different groups of students. The purpose of this research was to examine the development of oral reading fluency in a sample of second-grade students. We illustrated and compared growth models for groups of diverse students at different levels of risk.

Participants

The children were second graders who were participating in a research project focused on preventing behavior and reading problems in 14 elementary schools enrolling large numbers of at-risk students in a large urban/suburban school district in the southeastern region of the United States. The number of students from each school ranged from 35 to 152. We selected schools randomly from a pool meeting criteria similar to those used in other large-scale reading projects addressing the needs of at-risk students (e.g., Texas Reading Initiative):

- Recognized performance accountability ratings ("at standard" or above) on national, state, or local assessments.
- High rates of participation in the federal free and reduced-price lunch program.
- Evidence of effective implementation of the district's early reading intervention program and full allocation (at least 120 min) of literacy block instructional time.
- Willingness to use project-identified measures as evidence of reading and behavior improvement.
- Support of the Superintendent, Senior Staff, Principal, and campus site-based decision-making team, faculty, and staff.
- Willingness to serve as a demonstration site and to collaborate in efforts to mentor other schools.

The participants ($N = 1153$) included 668 students with at-risk and 485 (42%) with benchmark reading performance. There were 555 (48.1%) girls and 598 (51.9%) boys, and the participants were predominantly African American ($n = 662$, 57%) with 272 (24%) in Hispanic, 124 (11%) in European American, and 95 (8%) in Asian or other ethnic groups. There was a statistically significant relationship ($\chi^2 = 41.76$, $df = 3$) between ethnicity and reading status, with more Asian and European American students in the benchmark reading group and more Hispanic students in the at-risk group than expected by chance. Similarly, there was a statistically significant relationship ($\chi^2 = 10.82$, $df = 1$) between gender and reading status, with more girls in the benchmark reading group and more boys in the

at-risk group than expected by chance. The relationship between gender and ethnicity was not statistically significant ($\chi^2 = 1.81$, $df = 1$).

Core Reading Program[1]

Participating schools used the *Open Court Reading Program* (Open Court: SRA/McGraw Hill, 2000). In our study, teachers used the program materials during a daily district-wide 90-min literacy block of whole-class instruction followed by 30 min of small group instruction and/or independent work following guidelines presented in a synthesis of research on learning to read (Adams, 1990) and a large-scale research study by Foorman, Francis, Fletcher, Schatschneider, and Mehta (1998).

In the early stages of fluency development in Open Court, students read manageable texts that allow them to practice their growing knowledge of sound–symbol relationships. Typically, this means reading stories that contain a high number of words that can be "sounded out" based on what the student has been taught. Repeated practice reading words that used newly learned sounds and spellings helped students build fluency and vocabulary. Recognizing that accurate, rapid, expressive oral reading develops over time, Open Court has decodable texts at multiple grade levels to allow ample opportunities for practice and feedback.

In a recent evaluation, Skindrud and Gersten (2006) cited SRA materials and reported that the Open Court program "provides systematic instruction in the areas of decoding, comprehension, inquiry, investigation, writing, spelling, vocabulary, grammar, usage, mechanics, penmanship, listening, and speaking" (p. 393). They found that students who used Open Court outperformed their peers who used another program on reading and language scores on the Stanford Achievement Test, Ninth Edition (SAT9). They did not mention fluency in their description of the program or include indicators of it in their outcome measures.

Measurement

We were interested in relationships among oral reading rates and risk factors identified by previous research, such as gender, reading status, and ethnicity. Widely used and accepted standards served as measurement indicators.

[1]We were not evaluating the core reading program in the district. We provide a brief description of it as a basis for understanding and knowing the context in which our research took place.

Oral reading rates. The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) is a comprehensive system developed to assess essential beginning reading skills (Good & Kaminski, 2002, 2003). The oral reading fluency (ORF) measure tests accuracy and fluency in connected text against benchmark goals of 40 correct words per min by the end of first grade, 90 correct words per min by the end of second grade, and 110 words per min by the end of third grade. We administered second-grade passages following developers' guidelines at recommended times during the school year.[2]

In general, the DIBELS assessments have excellent technical adequacy (cf. Coyne & Harn, 2006; Elliott, Lee, & Tollefson, 2001; Fuchs, Fuchs, & Compton, 2004; Good, Kaminski, Simmons, & Kame'enui, 2001; Hintze, Ryan, & Stoner, 2003; Kaminski & Good, 1996; Speece, Mills, Ritchey, & Hill, 2003; Vadasy, Sanders, & Peyton, 2005). A series of studies have confirmed the technical adequacy of ORF procedures in general. For example, test–retest reliabilities for elementary students ranged from .92 to .97; alternate-form reliability of different reading passages drawn from the same level ranged from .89 to .94 (Tindal, Marston, & Deno, 1983). Criterion-related validity studied in eight separate studies in the 1980s reported coefficients ranging from .52 to .91 (Good & Jefferson, 1998).

Our data were grounded in best practices recommended by the test developers and used by the district for all other large-scale assessments (e.g., End-of-Grade accountability testing). Prior to the district-wide adoption of DIBELS, teachers received instruction in administering and scoring it from trained professional development providers and had opportunities to practice in their schools before assessing students in their classrooms for the first time. Literacy facilitators provided support at each school and entered individual student scores into the web-based database maintained by the University of Oregon. The tests were not used as any kind of accountability marker in the district. As with other similar studies, we have no information on the accuracy of data collection (cf. Schilling et al., 2007). Although this represents a potential limitation related to the reliability of the scores, having followed best practice recommendations provided by the DIBELS developers supervised by district-level accountability professionals, we reasoned that the technical accuracy of the extant scores was acceptable for the purposes of our research. In a similar large-scale assessment (i.e., the NAEP 2002 Special Study of Oral Reading, cited in Daane et al., 2005), trained scorers rated digital recordings of 1,779 fourth graders reading a passage aloud. In the report of that work, no indication was provided of the number

[2]See http://dibels.uoregon.edu

of judges, levels of interrater reliability, or the potential limitations that the procedure contributed to the outcomes.

We used initial scores to group students with respect to their initial oral reading proficiency levels using standards (at risk vs. benchmark) recommended for the DIBELS assessments.[3] According to the first of three assessments, 668 (57.9%) were considered at risk, with scores below 44 words per min; the rest of children (42.1%) were reading at benchmark (ORF $\geq$ 44). Schilling et al. (2007) found that these benchmarks were "reasonably accurate in identifying students whose reading was below average" on end-of-grade standardized achievement tests (p. 442). We were interested in describing and comparing differences in oral reading, as well as in predicting relationships between oral reading across reading group, gender, and ethnicity in a large sample of students at risk for continued school failure. Although data on oral reading rates are available (cf. Hasbrouck & Tindal, 1992, 2006), students in our research represent groups for whom more information and more detailed analyses are needed.

Design and Data Analysis

Oral reading was measured three times (fall, winter, and spring) during the second grade. We used analysis of variance with repeated measures (ANOVR) to examine changes in performance over time across gender, ethnicity, and reading status groups. Significant effects were followed with univariate analyses "as a preliminary step in a hierarchical data analysis" (Raudenbush & Bryk, 2002, p. 24). We used growth curve analysis to examine students' growth rate in oral reading, as well as to evaluate the extent to which individual characteristics influenced growth rates using Hierarchical Linear Modeling (HLM6; Raudenbush, Bryk, Cheong, Congdon, & Toit, 2004). Data were aggregated by the three time intervals across the variables of *gender*, *reading*, and *ethnicity*. Measurement intervals were recoded so that the intercept at Level 1 represented the initial reading proficiency of individual students at the beginning of the second grade (Raudenbush & Bryk, 2002). We built two models to address our research questions.

First, unconditional models (without predictors) were used to examine the mean and variance of the within-subject parameters. Unconditional models were fit first to provide useful empirical evidence for determining a proper specification of the individual growth equation and baseline statistics for evaluating subsequent conditional models (Raudenbush & Bryk, 2002). Two parameters were of interest in describing the unconditional

[3]See http://dibels.uoregon.edu/benchmarkgoals.pdf

modes: intercept and slope. The intercept represented the initial status of a student's reading proficiency and the slope represented the child's growth rate during the second grade. Model parameters were tested sequentially (intercept and slope), first examining the fixed effect and then the random effect.

In the second set of analyses, we were interested in examining the differences in means and growth rates between at-risk students and benchmark students, as well as between boys and girls and across ethnic groups. Therefore, a two-level conditional model was tested. The first level estimated the overall change of all students' oral reading with time, the second level assessed the impact of the children's characteristics (gender, reading status, and ethnicity) on their initial oral reading scores and the rate of the change in oral reading. Conditional models (with predictors) were used to estimate the cause of the variance of the within-subject parameters. In this way, we could examine differences in levels and growth parameters between various groups of students defined by their characteristics.

RESULTS

We described and compared students' performance on ORF as a function of time across reading status, gender, and ethnicity. The means and standard deviations for fall, winter, and spring assessments across reading status, gender, and ethnicity are in Table 1. We compared these scores with an ANOVR procedure with reading status, gender, and ethnicity as the between-subjects factor and time (beginning, middle, and end of the school year) as the within-subjects factor. The assumption of homogeneity of variance was not met; Box's $M = 588.48$, $F\,(90, 101367.20) = 6.33$, $p < .001$; and because the sample sizes for each ethnic group were quite different from each other, we used a stringent alpha level ($\alpha = .01$) instead of transforming scores to reduce variability because of the concern for the limitation of interpretation (Tabachnick & Fidell, 2007).

For tests of within-subject effects, no significance was noted for any of the two-way or three-way interactions but a significant main effect for time was obtained, $F\ (2,\ 2274) = 1822.55$. Oral reading improved from fall ($M = 49.37$, $SD = 31.17$) to winter ($M = 73.84$, $SD = 36.57$) to spring ($M = 89.77$, $SD = 37.75$) assessments. Because time is a quantitative variable equal in interval, trend analysis was justified. A significant linear trend across time was noted; $F\,(1, 1137) = 2508.73$, $\eta^2 = .69$. A quadratic trend was also statistically significant; $F\,(1, 1137) = 98.52$, $\eta^2 = .08$.

For tests of between-subject effects, none of the two-way and three-way interactions were significant. All the main effects for gender, ethnicity, and reading status were significant, $F\ (1,\ 1137) = 11.19$, $\eta^2 = .01$ for gender;

TABLE 1
Oral Reading Performance across Reading, Gender, and Ethnicity Groups

Group			*Benchmark*		
Reading	*Gender*	*Ethnicity (n)*	*Fall*	*Winter*	*Spring*
At risk	Female	Asian	35.07 (10.66)	58.33 (19.55)	77.73 (26.39)
		African American	31.23 (11.60)	55.94 (21.07)	72.59 (26.90)
		Hispanic	23.26 (14.23)	47.28 (26.02)	64.92 (28.48)
		Caucasian	32.11 (9.97)	55.48 (19.55)	72.15 (20.12)
	Male	Asian	27.65 (14.56)	50.25 (22.45)	69.70 (26.86)
		African American	28.50 (12.18)	49.72 (21.88)	65.78 (25.88)
		Hispanic	23.33 (13.63)	42.95 (24.00)	62.50 (31.88)
		Caucasian	30.84 (13.39)	54.69 (27.27)	68.50 (28.07)
No risk	Female	Asian	91.13 (26.04)	116.67 (22.60)	128.37 (21.04)
		African American	77.19 (20.81)	106.66 (22.70)	120.43 (25.85)
		Hispanic	74.07 (16.57)	106.73 (21.96)	120.44 (24.36)
		Caucasian	96.85 (36.26)	119.78 (36.46)	131.42 (31.21)
	Male	Asian	74.17 (23.20)	97.47 (24.07)	114.90 (24.57)
		African American	74.74 (22.45)	98.86 (22.93)	113.85 (23.77)
		Hispanic	70.82 (14.06)	100.84 (19.92)	115.42 (26.48)
		Caucasian	92.47 (32.17)	117.82 (37.11)	135.42 (37.29)

F (1, 1137) = 1031.06, $\eta^2 = .48$ for reading status; and F (3, 1137) = 10.91, $\eta^2 = .03$ for ethnicity. A Scheffé test was used as the post hoc test because it is one of the safest of all possible post hoc tests and does not require equal sample sizes for all groups (Gravetter & Wallnau, 2007). Overall performance of benchmark students ($M = 101.72$, $SD = 23.50$) was statistically higher than that of at-risk students ($M = 48.69$, $SD = 19.92$). Overall performance of girls ($M = 75.81$, $SD = 33.93$) was statistically higher than that of boys ($M = 66.53$, $SD = 33.24$). The post hoc test revealed that Caucasian ($M = 85.45$, $SD = 41.97$) students performed significantly better than the African American ($M = 70.91$, $SD = 31.29$) group, and the African American group performed significantly better than the Hispanic ($M = 59.76$, $SD = 32.08$) group. The difference between Caucasian and Asian/Other ($M = 84.91$, $SD = 32.76$) students was not statistically significant.

Due to stratified random selection method, our data have three levels—school, teacher, and students—in which students are nested within classrooms (teacher), and classrooms are nested within schools. Intraclass correlation coefficient (ICC), a ratio of between-group variance over the total variance, was used to examine the percentage of variance explained by each level of teacher and school (Stapleton, 2006).

Where

$$ICC = \frac{MS_B - MS_W}{MS_B + (n-1)MS_W}$$

For the teacher level, students were clustered within classrooms but the clusters do not contain equal numbers of participants. Therefore, the sample size per group if the clusters are balanced was calculated using the following formula provided by Kenny and Judd (1986).

$$n = \frac{N^2 - \sum_{g=1}^{G} n_g^2}{N(G-1)} = \frac{1153^2 - \sum_{g=1}^{16} n_g^2}{1153(72-1)} \approx 15.99$$

The average cluster size for the whole sample was 16.39 with a standard deviation of 4.46 (min = 2; max = 23). So, the calculated sample size per group if the sample size were balanced across groups (n) was quite close to the average cluster size. Components of ANOVR using school and teacher as the grouping variable, and the resulting estimates of the ICC for each of the three variables are in Table 2. These values were all small in size for both the teacher level and the school level (Kreft & de Leeuw, 1998), indicating that around 6% of the variance in each of the variables can be attributed to school effects and 21% of the variance can be attributed to teacher effects. Most of the variance (nearly 80%) is attributed to individual effects. This conclusion was confirmed with a three-level HLM (the between-school variance is only 5.94% of the total variance). As a result, a two-level HLM was employed with the first level representing the change of individuals' ORF scores across time, and the second level representing

TABLE 2
ANOVR Components and Estimates of Intraclass Correlation Coefficients (ICC) for Each Variable

	MS_W	MS_B	ICC
Teacher level			
ORF (Fall)	770.99	4022.98	.21
ORF (Winter)	1101.09	4930.70	.18
ORF (Spring)	1088.09	6559.75	.24
School level			
ORF (Fall)	908.97	6443.07	.07
ORF (Winter)	1284.36	5959.11	.04
ORF (Spring)	1330.41	9741.03	.07

Note. MS_w = Mean Square within. MS_B = Mean Square between. ORF = Oral Reading Fluency.

individual differences with respect to gender, reading status, and ethnicity. Each ethnic group was dummy coded (0 vs. 1) for model building except Asian/Other, because no differences were indicated between students in this group and their Caucasian peers.

Unconditional Model

The developmental trends of oral reading for all participating second-grade students are illustrated in Figure 1 for gender, Figure 2 for reading groups, and Figure 3 for ethnicity. A consistent achievement gap is evident across the three benchmark assessments. The unconditional model of oral reading was a random-intercept and random-slope model (see Table 3). The intercept represents the status of each child at the beginning of second grade, and the slope is the instantaneous growth rate for each child each month. The quadratic term captures the curvature or acceleration in each growth trajectory. Time was coded as 0, 5, and 10, which represents the beginning (fall) of second grade, 5 months later (winter), and 10 months later (spring). On average, the predicated fall ORF score was 49.37 words per min and the predicted monthly growth was 5.75 words per min. The curvature in each growth trajectory is -0.17 words per min, suggesting a decreasing growth rate in the second grade. For example, the average monthly growth rate for a child in Winter is 4.05 words per min $(5.75 + 2 \times (-0.17) \times 5)$ and in Spring is 2.35 words per min $(5.75 + 2 \times (-0.17) \times 10)$. The correlation

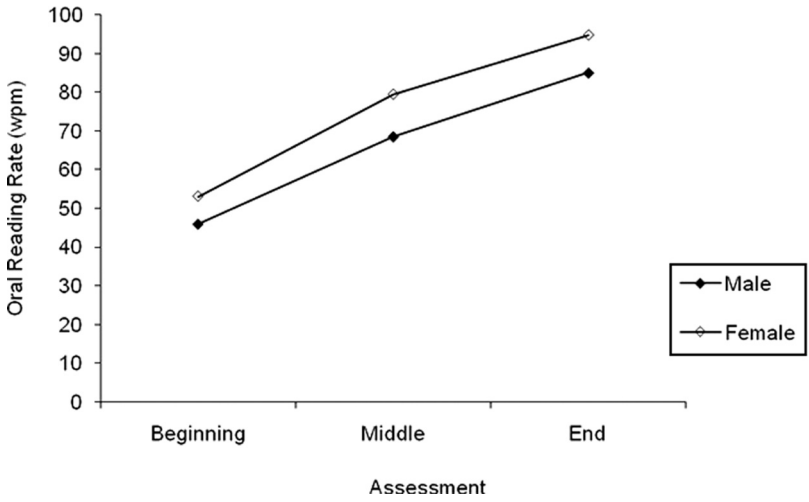

FIGURE 1 Second grade developmental oral reading trend by gender.

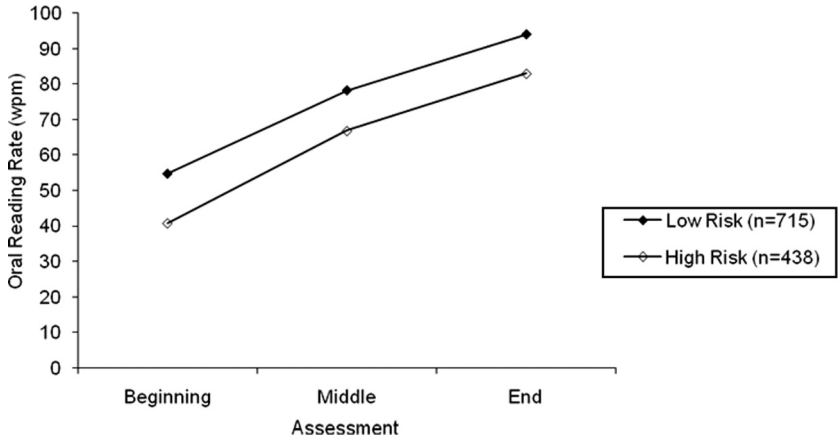

FIGURE 2 Second grade developmental oral reading trend by reading status.

between the initial status of oral reading (intercept) and the learning rate (slope) during the second grade was estimated to be .22. The reliability for the parameters (intercept and slope) were .93 and .59, respectively.

As an exploratory analysis, unconditional models of the student-level coefficients and their standard errors were obtained by regressing Empirical

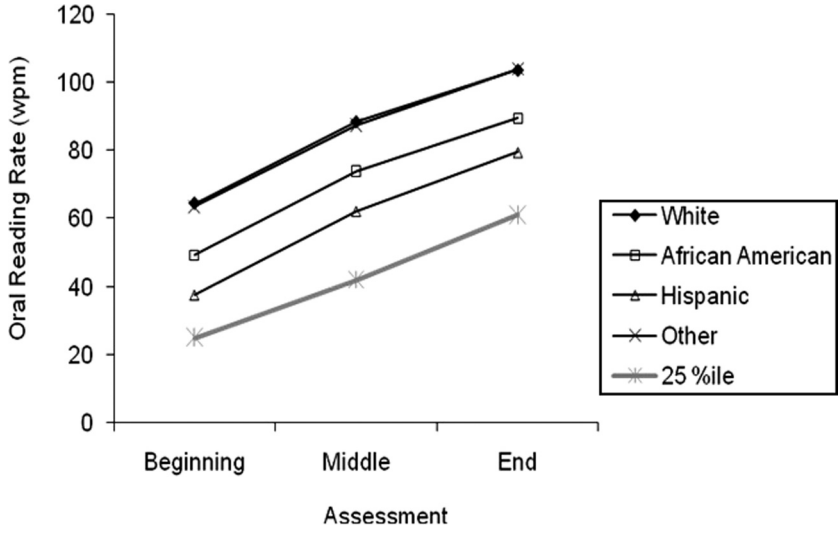

FIGURE 3 Second grade developmental oral reading trend by ethnicity.

TABLE 3
Three-Level Analyses of Oral Reading Fluency During the Second Grade
(Unconditional Model at Levels 2 and 3)

Fixed effect	*Coefficient*	*SE*	*Observed* t	*p*
Initial status (intercept)	49.37	0.92	53.82	<.001
Linear growth rate (slope)	5.75	0.15	39.23	<.001
Quadratic growth rate (curvature)	−0.17	0.01	− 12.88	<.001
Random effect	*Variance*	*df*	χ^2	*p*
Initial status (intercept)	951.94	1152	16736.48	<.001
Linear growth Rate (slope)	2.43	1152	2809.31	<.001

Bayes residuals on student predictors selected for possible inclusion in subsequent conditional models. The student variables, gender ($t = -4.51$), reading status ($t = 44.97$), Hispanic or not ($t = -6.91$), and Caucasian or not ($t = 5.43$), were selected for the intercept but only gender ($t = -3.15$) and reading status ($t = 8.37$) were included for the slope for subsequent conditional models. Other variables were not included due to small absolute t values (less than 1).

Conditional Model

The correlation between the initial status of oral reading and the learning rate during the second grade was estimated to be .32. The reliability for the parameters was .80 and .59, respectively. Final estimates of the fixed and random effects are in Table 4. The initial mean oral reading score for all students was 30.28 words per min. Male students had significantly lower initial mean oral reading scores than female students, $t(1148) = -3.16$. Benchmark students had significantly higher initial mean oral reading scores than at-risk students, $t(1148) = 43.67$, Hispanic students had significantly lower initial mean oral reading scores than students of other ethnic groups, $t(1148) = -5.27$, and Caucasian students had significantly higher initial mean oral reading scores than students of other ethnic groups, $t(1148) = 3.83$.

The mean monthly growth rate for all students was significantly positive at 5.80 words per min, $t(1150) = 33.57$. Male students' monthly growth rate was slower than that of female students, $t(1150) = -2.07$, but no significant differences were noted for the monthly growth rates between benchmark and at-risk students. The comparison between the conditional and unconditional models revealed that the conditional model explained 71.22% of the variance in initial oral reading scores for second-grade students but only

TABLE 4
Effects of Children's Characteristics on Oral Reading Fluency

Fixed effect	*Coefficient*	*SE*	*Observed* t	*p*
Intercept				
Initial Status	30.28	0.88	34.48	<.001
Gender	−3.46	1.10	−3.16	.002
Reading Status	50.48	1.16	43.67	<.001
Hispanic	−5.84	1.11	−5.27	<.001
White	9.63	2.52	3.83	<.001
Slope				
Gender	−0.25	0.12	−2.07	.039
Reading Status	0.18	0.12	1.51	.131

Random effect	*Variance*	*df*	χ^2	*p*
Initial Status (intercept)	273.94	1148	5616.55	<.001
Linear Growth Rate (slope)	2.41	1150	2792.36	<.001

0.82% of the variance of the monthly growth rate. This conclusion is consistent with that from ANOVR, suggesting large differences between gender, reading status, and ethnicity with respect to reading skills but no interaction between either of these variables with time. The reading skills gap between these groups of students existed at the beginning of the second grade and remained nearly the same throughout the second grade.

DISCUSSION

Despite sustained interest from professionals, fluency, or "the ability to read with prosody," was "rarely found in state blueprints and test specifications in a recent analysis of constructs assessment in high stakes assessments" (Johnstone et al., 2007, p. 5) and few studies have documented or compared oral reading rates or development. We evaluated the reading rates in schools enrolling large numbers of students at risk for school failure. Hasbrouck and Tindal (1992, 2006) provided ORF percentile norms for students in Grades 2–5 and 1–8, respectively. In their work, they did not differentiate scores for boys and girls or for students from diverse ethnic groups. They also did not provide information about the reading program(s) used in the schools included in their studies. Our research replicates and expands the knowledge base on ORF.

Children in our study participated in a district-wide, evidence-based core reading program (i.e., Open Court). Their teachers received extensive

professional development and regular classroom-based support. They used pacing guides and were expected to follow scripted lessons. We did not evaluate the benefits of implementing Open Court in efforts to improve reading achievement in a large school district; in fact, fluency instruction is not a big part of this program (cf. Skindrud & Gersten, 2006). We were interested in the development of and variation in oral reading reflected in extant scores gathered by district personnel. Unlike in previous reports of oral reading performance, the program in which students participated was known and consistent across the schools they attended. We believe the application of the same core reading program is a strength of our study (i.e., development and variation are less likely due to differences in programs provided or vagaries of implementing them across schools, classrooms, and teachers).

Based on the extant literature, we expected differences in oral reading across groups of children reading at different initial levels. In schools providing evidence-based core reading instruction, the rich (benchmark readers) get richer and the poor (students at risk) progress, but remain below their higher achieving peers. This finding replicates what was known about oral reading rates in second grade based on data provided by Hasbrouck and Tindal (1992, see Figure 4). The outcome speaks to the need for using multilevel models (cf., Vaughn & Linan-Thompson, 2003) "to aid in improving the reading skills of students who are poor readers" (Stewart, Benner,

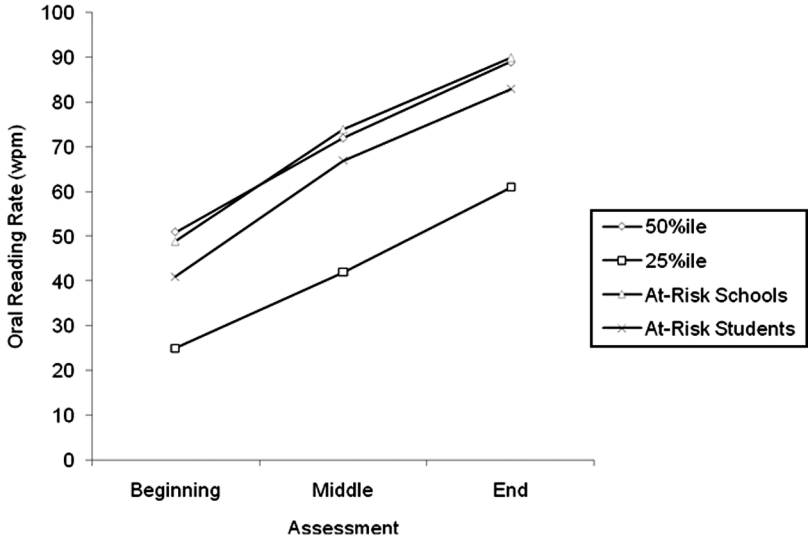

FIGURE 4 Second grade developmental oral reading trend across studies.

Martella, & Marchand-Martella, 2007, p. 239). In these approaches, all students participate in core reading instruction, and supplemental reading instruction is provided to students who are not making progress in the primary program. Some estimates place the number of these students at 30% to 40% of the school population (cf. Coyne, Kame'enui, & Simmons, 2001); almost 60% of the students in schools enrolling large numbers of at-risk students such as those in our study would qualify for supplemental instruction based on their beginning oral reading rates. Without attention, the trajectory of failure for these children is unchanged and the gap between their performance and that of their peers remains formidable.

The overall reading achievement scores of 9-year-old girls have been consistently higher than those of same-age boys since the early 1970s, when the National Assessment of Educational Progress (NAEP) began documenting the mastery of basic skills in America's youth (cf. U. S. Department of Education, 2006, Table 108). Scores on these measures are reported on a scale ranging from 0 to 500. NAEP scores are not indicators of oral reading rates or fluency. Students at reading score level 200 are able to understand, combine ideas, and make inferences based on short uncomplicated passages about specific or sequentially related information. Students at reading score level 250 are able to search for specific information, interrelate ideas, and make generalizations about literature, science, and social studies materials. Students at reading score level 300 are able to find, understand, summarize, and explain relatively complicated literary and informational material. Scores reported for 9-year-old boys (201–216) were consistently 5–13 points below scores for girls (214–221) in data available from 1971 to 2004 (U. S. Department of Education, 2006, Table 108). Daane et al. (2005) found that fourth grade girls outperformed boys on three measures of oral reading (i.e., accuracy, rate, and fluency). Hasbrouck and Tindal (1992) did not disaggregate their data by gender. None of the previously reported research reflects oral reading over time (e.g., across the school year). In our study, oral reading rates of girls were consistently higher than those of boys across three assessments in second grade.

For at least 30 years, research has documented that poor students of color and those with academic problems are also likely to be pushed out of school through exclusionary discipline practices with the deleterious, albeit unintended consequence of reduced opportunities to learn (Children's Defense Fund, 1975; Fenning & Rose, 2007). We did not observe children's behavior or discipline events in the participating at-risk schools. We did find consistent gaps in the oral reading performance of students of color who began the school year at risk of continued school failure compared to their peers. National trends in overall reading achievement for 9-year-olds reported from 1975–2004 are consistent for different ethnic groups; on

average, Caucasian students scored 20–40 points higher than their African American and Hispanic peers (U. S. Department of Education, 2006, Table 108). Daane et al. (2005) documented differences in oral reading across gender and ethnic groups. Only a portion of these national outcomes reflect oral reading accuracy, rate, and fluency; other extant oral reading norms are not reported by ethnicity. In our study, oral reading rates were consistently higher for Caucasian and Asian students than for their African American and Hispanic peers.

In the absence of any other intervention, children in low reading groups made progress in an evidence-based core reading program but remained below their classmates who began the year with higher levels of oral reading performance. Data from the NAEP *Special Study of Oral Reading* (Daane et al., 2005) have suggested that these differences persist across gender and ethnic groups at least into fourth grade. Again, our work points to the need for prevention models aimed at catching students before they fall behind and providing support using alternative instructional models.

Implications for Future Research and Practice

Our work focused on the development of oral reading in at-risk students over the second-grade school year. We were interested in describing and comparing differences in oral reading using test-developer-recommended benchmarks, as well as in predicting relationships between oral reading across reading group, gender, and ethnicity in a large sample of students at risk for continued school failure. Our study was not an evaluation of efforts to teach fluency, but more a record of its development within the context known to us as a result of a core reading program being used across the district. Studies of the effects of implementing intensive oral reading interventions across reading groups, gender, and ethnicity are clearly warranted.

Schools, districts, and states seeking to bring additional resources (i.e., more intensive and explicit instruction) to struggling readers would surely benefit from the identification of those students who are not performing and not showing adequate gains in oral reading. Using data to make decisions about ongoing student performance and to plan instruction (i.e., progress monitoring) is an evidence-based practice with tremendous promise for improving the lives and academic futures of all children, not just those at risk of academic failure:

> When teachers use systematic progress monitoring to track their students' progress in reading, mathematics, or spelling, they are better able to identify students in need of additional or different forms of instruction, they design stronger instructional programs, and their students achieve better. (Fuchs & Fuchs, 2002, p. 1; also cited in Safer & Fleischman, 2005, p. 81)

In our study, we did not study progress monitoring. Similarly, we did not study the predictive value of DIBELS ORF performance. Schilling et al. (2007) found that the same benchmarks used in our research were "reasonably accurate in identifying students whose reading was below average" on end-of-grade standardized achievement tests (p. 442). Clearly, studies of the usefulness of these practices and markers, as well as other "cut" scores and performance indicators in monitoring progress and predicting outcomes on high-stakes tests, are needed.

Further, although fluent reading does not ensure understanding (cf. Jenkins et al., 2003; Yovanoff, Duesberry, Alonzo, & Tindal, 2005), links between oral reading rates and comprehension have been demonstrated (cf. Daane et al., 2005; National Reading Panel, 2000; Schilling et al., 2007). The consistent differences in oral reading rates for boys and girls, children from different ethnic backgrounds, and students with different initial oral reading rates that we observed provoke us to seek explanations for the consistent variance observed over the years in NAEP and other high-stakes testing reading scores. Continued research comparing relationships between oral reading and comprehension across groups is warranted.

Current best practice in education is driven by including, representing, and disaggregating the performance of all students in local, state, and national reporting of performance on large-scale assessments (Fast, Blank, Potts, & Williams, 2002; No Child Left Behind, 2002; Thurlow & Wiley, 2006). Public reporting of educational progress supports informed decision making about the effectiveness of schools and their programs. It also provides a basis for evaluating how well students are doing in reform contexts and addressing important issues such as the achievement gap between groups of students (Center on Education Policy, 2004; Thurlow & Wiley, 2006). The clear differences in oral reading rates that we observed reaffirm the importance of disaggregation of scores when documenting the performance of students, especially those at risk for continued school failure. The differences that we observed were evident at each benchmark assessment. The value of disaggregated reporting for early identification and intervention is obvious in efforts to close the gaps between children from diverse backgrounds and their peers.

Extant "norms" for ORF do not reflect similarities or differences within and between children attending U.S. schools. Future research replicating our findings across schools using different core reading programs in different geographic regions of the country will enrich the extant and developing knowledge base. Similarly, there is value in illustrating and comparing oral reading and other early literacy skill performance across the diverse groups of children represented in the classrooms of most schools. The progress that

may come from such efforts is promising; the perils of not engaging in them are myriad.

ACKNOWLEDGMENTS

Support for this research was provided in part by Grant No. H238X00001 from the U.S. Department of Education, Office of Special Education Programs, awarded to the University of North Carolina at Charlotte. The opinions expressed do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement should be inferred.

REFERENCES

Abadiano, H. R., & Turner, J. (2005). Reading fluency: The road to developing efficient and effective readers. *New England Reading Association Journal*, *41*(1), 50–56.

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Behavioral Research and Teaching. (2004). *Oral reading fluency: 90 years of measurement* (Technical Report #33). Eugene: University of Oregon.

Cassidy, J., & Cassidy, D. (2003). What's hot, what's not for 2004. *Reading Today*, *21*(5), 1, 18.

Cassidy, J., & Cassidy, D. (2004a). Literacy trends and issues today: An on-going study. *Reading and Writing Quarterly*, *20*, 11–28.

Cassidy, J., & Cassidy, D. (2004b). What's hot, what's not for 2005. *Reading Today*, *22*(3), 1.

Cassidy, J., & Cassidy, D. (2005). What's hot, what's not for 2006. *Reading Today*, *23*(3), 1.

Cassidy, J., & Cassidy, D. (2007). What's hot, what's not for 2007. *Reading Today*, *24*(4), 1.

Cassidy, J., Garrett, S., & Barrera IV, E. S. (2006, September). What's hot in adolescent literacy 1997–2006. *Journal of Adolescent & Adult Literacy*, *50*, 30–36.

Center on Education Policy. (2004). *From the capital to the classroom: Year 2 of the No Child Left Behind Act*. Washington, DC: Author.

Chard, D. J., Pikulski, J. J., & McDonagh, S. H. (2006). Fluency: The link between decoding and comprehension for struggling readers. In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 39–61). New York: Guildford.

Chatel, R. G. (2005). Developing reading fluency: An abundance of technology resources. *New England Reading Association Journal*, *41*(1), 57–60.

Children's Defense Fund. (1975). *School suspensions: Are they helping children?* Cambridge, MA: Washington Research Project.

Coyne, M. D., & Harn, B. A. (2006). Promoting beginning reading success through meaningful assessment of early literacy skills. *Psychology in the Schools*, *43*, 33–43.

Coyne, M. D., Kame'enui, E. J., & Simmons, D. C. (2001). Prevention and intervention in beginning reading: Two complex systems. *Learning Disabilities Research & Practice*, *16*, 62–73.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*, 934–945.

Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Oranje, A. (2005). *Fourth-grade students reading aloud: NAEP 2002 Special study of oral reading* (NCES 2006-469). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, Government Printing Office.

Dudley, A. M., & Mather, N. (2005). Getting up to speed on reading fluency. *New England Reading Association Journal*, *41*(1), 22–27.

Ehri, L. C., Nunes, S. R., Stahl, S., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393–447.

Ellery, V. (2004). *Creating strategic readers: Techniques for developing competency in phonemic awareness, phonics, fluency, vocabulary, and comprehension*. Newark, DE: International Reading Association.

Elliott, J., Lee, S. W., & Tollefson, N. (2001). A reliability and validity study of the Dynamic Indicators of Basic Early Literacy Skills-Modified. *School Psychology Review*, *30*, 33–49.

Fast, E. F., Blank, R. K., Potts, A., & Williams, A. (2002). *A guide to effective accountability reporting*. Washington, DC: Council of Chief State School Officers.

Fenning, P., & Rose, J. (2007). Overrepresentation of African American students in exclusionary discipline: The role of school policy. *Urban Education*, *42*, 536–559.

Foorman, B. R., & Moats, L. C. (2004). Conditions for sustaining research-based practices in early reading instruction. *Remedial and Special Education*, *25*, 51–60.

Foorman, B. R., & Torgeson, J. (2001). Critical elements of classroom and small group instruction promote reading success in all children. *Learning Disabilities Research and Practice*, *16*, 203–212.

Foorman, B., Francis, D., Fletcher, J., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, *90*, 37–55.

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. (1996). Combining performance assessment and curriculum-based measurement to strengthen instructional planning. *Learning Disabilities Research and Practice*, *11*, 183–192.

Fuchs, L. S., & Fuchs, D. (2002). *What is scientifically-based research on progress monitoring?* (Technical report). Nashville, TN: Vanderbilt University.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41*, 93–99.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification versus nonsense word fluency. *Exceptional Children*, *71*, 7–21.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Computers and curriculum-based measurement: Effects of teacher feedback systems. *School Psychology Review*, *18*, 112–125.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, *9*, 20–28.

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

Good, R. H., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 61–88). New York: Guilford.

Good, R. H., & Kaminski, R. A. (2002). *DIBELS oral reading fluency passages for first through third grades* (Technical Report No. 10). Eugene, OR: University of Oregon.

Good, R. H., & Kaminski, R. A. (2003). *Administration and scoring guide: Dynamic Indicators of Basic Early Literacy Skills* (DIBELS). Longmont, CO: Sopris West.

Good, R. H., III, Kaminski, R. A., Simmons, D., & Kame'enui, E. J. (2001). *Using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an outcomes-driven model: Steps to reading outcomes* (unpublished manuscript). Eugene: University of Oregon at Eugene.

Good, R. H., Wallin, J. U., Simmons, D. C., Kame'enui, & Kaminski, R. A. (2002). *System-wide percentile ranks for DIBELS benchmark assessment* (Technical Report 9). Eugene, OR: University of Oregon.

Gravetter, F. J., & Wallnau, L. B. (2007). *Statistics for the behavioral sciences*. Belmont, CA: Wadsworth/Thomson Learning.

Hasbrouck, J. E., & 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. E., & Tindal, G. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, *59*, 636–644.

Hintze, J. M., Ryan, A. L., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the Dynamic Indicators of Basic Early Literacy Skills and the Comprehensive Test of Phonological Processing. *School Psychology Review*, *32*, 541–556.

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

Johns, J. L., & Berglund, R. L. (2005). *Fluency: Questions, answers, evidence-based strategies* (2nd ed.). Dubuque, IA: Kendall-Hunt.

Johnstone, C. J., Moen, R. E., Thurlow, M. L., Matchett, D., Hausmann, K. E., & Scullin, S. (2007). *What do state reading test specifications specify*? Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Kame'enui, E. J., & Simmons, D. C. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading*, *5*, 203–210.

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

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, *99*, 422–431.

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling.* Newbury Park, CA: Sage.

Kuhn, M. (2004). Helping students become accurate, expressive readers: Fluency instruction for small groups. *The Reading Teacher*, *58*, 338–344.

Kuhn, M. R., & Stahl, S. A. (2000). *Fluency: A review of developmental and remedial practices* (Center for the Improvement of Early Reading Achievement Report No. 2–008). Ann Arbor, MI: University of Michigan.

Kuhn, M. R., Schwanenflugel, P. J., Morris, R. D., Mandel Morrow, L., Gee, W., Meisinger, E. B., et al. (2006). Teaching children to become fluent and automatic readers. *Journal of Literacy Research*, *38*, 357–387.

Madelaine, A., & Wheldall, K. (2004). Curriculum-based measurement of reading: Recent advances. *International Journal of Disability, Development, and Education*, *51*, 57–82

Marsell, B. (2007). Fluent to a fault: Put fluency in the passenger seat and let comprehension take the wheel! *Reading Today*, *24*(6), 18.

Mesmer, H. A. E. (2006). Beginning reading materials: A national survey of primary teachers' reported uses and beliefs. *Journal of Literacy Research*, *38*, 389–425.

Moats, L. C. (2001). When older students can't read. *Educational Leadership*, *58*, 36–40.

National Institute of Child Health and Human Development. (2000a). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific*

research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

National Institute of Child Health and Human Development. (2000b). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Retrieved July 31, 2003, from http://www.nichd.nih.gov/publications/nrp/smallbook.htm

National Institute of Child Health and Human Development. (2000c). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press. Retrieved December 1, 2005, from http://books.nap.edu/books/030906418X/html/index.html

No Child Left Behind Act of 2001, 20 U.S.C. § 6301 *et seq*. (2002). Retrieved November 1, 2007, from http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Oakley, G. (2005). Reading fluency as an outcome of a repertoire of interactive reading competencies: How to teach it to different types of dysfluent readers (and how ICT can help). *New England Reading Association Journal*, *41*(1), 13–21.

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *Reading Teacher*, *58*, 510–519.

Prescott-Griffin, M. L., & Witherell, N. L. (2004). *Fluency in focus: Comprehension strategies for all young readers*. Portsmouth, NH: Heinemann.

Rasinski, T. (2005). The role of the teacher in effective fluency instruction. *New England Reading Association Journal*, *41*(1), 9–12.

Rasinski, T. V. (2000). Speed does matter in reading. *The Reading Teacher*, *54*, 146–151.

Rasinski, T. V. (2003). *The fluent reader: Oral reading strategies for building word recognition, fluency, and comprehension*. New York: Scholastic.

Rasinski, T., Blachowicz, C. L., & Lems, K. (2006). Introduction. In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 1–3). New York: Guildford.

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

Raudenbush, S., Bryk, A., Cheong, Y. F., Congdon, R., & Toit, M. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Reutzel, D. R. (2006). "Hey teacher, when you say 'fluency,' what do you mean?" In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 63–85). New York: Guildford.

Safer, N., & Fleischman, S. (2005). How student progress monitoring improves instruction. *Educational Leadership*, *62*(5), 81–83.

Samuels, S. J. (1997). The method of repeated readings. *The Reading Teacher*, *50*, 376–381.

Samuels, S. J. (2006). Reading fluency: Its past, present, and future. In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 7–20). New York: Guildford.

Samuels, S. J., Ediger, K., & Fautsch-Patridge, T. (2005). The importance of fluent reading. *New England Reading Association Journal*, *41*(1), 1–8.

Schilling, S. G., Carlisle, J. F., Scott, S. E., & Zeng, J. (2007). Are fluency measures accurate predictors of reading achievement? *Elementary School Journal*, *107*, 429–448.

Shanahan, T. (2006). Developing fluency in the context of effective literacy instruction. In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 21–38). New York: Guildford.

Skindrud, K., & Gersten, R. (2006). An evaluation of two contrasting approaches for improving reading achievement in a large urban district. *Elementary School Journal*, *106*, 389–407.

Speece, D. L., & Ritchey, K. D. (2005). A longitudinal student of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities*, *38*, 387–399.

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

SRA/McGraw Hill. (2000). *Open court reading* 2000 (2000 ed.). Retrieved on July 25, 2008, from https://www.sraonline.com/products.html?PHPSESSID=827ac331b8873410f2e9747d3965b226&tid=16&sid=1558&open=.

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*, 407–419.

Stage, S. A., Sheppard, J., Davidson, M. M., & Browning, M. M. (2001). Predicting of first-graders' growth in oral reading fluency kindergarten letter fluency. *Journal of School Psychology*, *39*, 225–237.

Stapleton, L. M. (2006). Using multilevel structural equation modeling techniques with complex sample data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 345–383). Charlotte, NC: Information Age.

Starch, D. (1915). The measurement of efficiency in reading. *The Journal of Educational Psychology*, *6*, 1–24.

Stewart, R. M., Benner, G. J., Martella, R. C., & Marchand-Martella, N. E. (2007). Three-tier models of reading and behavior: A research review. *Journal of Positive Behavior Interventions*, *9*, 239–253.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. New York: Pearson.

Thurlow, M. L., & Wiley, H. I. (2006). A baseline perspective on disability subgroup reporting. *The Journal of Special Education*, *39*, 246–254.

Tindal, G., Marston, D., & Deno, S. L. (1983). *The reliability of direct and repeated measurement* (Research Rep. 109). Minneapolis, MN: University of Minnesota Institute for Research on Learning Disabilities.

U. S. Department of Education. (2006). *Digest of Education Statistics, 2005* (NCES 2006-030). Washington, DC: National Center for Education Statistics.

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring: How text matters. *Journal of Learning Disabilities*, *38*, 364–380.

Vaughn, S., & Linan-Thompson, S. (2003). What is special about special education for students with learning disabilities? *Journal of Special Education*, *37*, 140–147.

Walker, B. J., Mokhtari, K., & Sargent, S. (2006). Reading fluency: More than fast and accurate reading. In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 86–105). New York: Guildford.

Wilson, S. L. (2005). Fluency in the classroom from read aloud to independent reading: Assessment, collaboration, practice, and performance. *New England Reading Association Journal*, *41*(1), 41–44.

Witte-Townsend, D. L., & Whiting, A. G. (2005). Seeking best practice in reading fluency development: Children, parents, and teachers tell stories of complex relations. *New England Reading Association Journal*, *41*(1), 28–34.

Yovanoff, P., Duesbery, L., Alonzo, J., & Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement: Issues and Practice*, *24*, 4–12.