Technical Adequacy Supplement for

Acadience™ Reading Oral Reading Fluency

Elizabeth N. Dewey
Kelly A. Powell-Smith, Ph. D.
Roland H Good III, Ph. D.
Ruth A. Kaminski, Ph. D.

Suggested Citation:

Dewey, E. N., Powell-Smith, K. A., Good, R. H., Kaminski, R. A., (2014) *Technical Adequacy Supplement for Acadience™ Reading Oral Reading Fluency.* Eugene, OR: Dynamic Measurement Group, Inc.

Technical Adequacy Supplement for Acadience™ Reading Oral Reading Fluency

The purpose of this document is to combine the various pieces of technical adequacy information for Acadience Reading Oral Reading Fluency (ORF) (also published under the DIBELS Next® mark) into one easily-accessible document. This document will serve multiple purposes: 1) to provide a quick reference to technical adequacy information, 2) to support the efforts of our research partners in submitting Acadience Reading for review to various agencies, and 3) to communicate new technical adequacy information that is available elsewhere, but has yet to be incorporated into the *Acadience Reading Technical Manual* (Good et al., 2013).

Similar information for other Acadience Reading measures is available in the *Acadience Reading Technical Manual* (Good et al., 2013), retrievable from Dynamic Measurement Group's website at https://acadiencelearning.org/.

Technical adequacy for Oral Reading Fluency (ORF) was evaluated through several studies. The technical reports for these studies are also available to download for free from DMG's website. A summary of the technical adequacy is reported at the end of this document in Table 1.

About the passages. The Oral Reading Fluency passages were selected from a group of passages evaluated during the *DIBELS Next Oral Reading Fluency Readability Study*[1] (Powell-Smith, Good, & Atkins, 2010). In this study, 20 individual progress monitoring passages were administered to students over a period of four to seven days (students read approximately two to five passages per day). All passages (i.e., both progress monitoring and screening) were designed to be equivalent, and all passages were evaluated for equivalency. The passages selected for screening (i.e., benchmark assessment) were chosen for their representativeness to all of the grade-level passages, and were sorted into groups of three, called an ORF Triad. The median score of the ORF Triad is the student's score. Those passages not selected for screening were selected for progress monitoring. Findings from the *DIBELS Next Oral Reading*

[1] Acadience Reading is also published under the name DIBELS Next. Some historical supporting documents contain the original name.

Readability Study (Powell-Smith, Good, & Atkins, 2010) show that the individual progress monitoring passages differ from each other about as much as the three passages within the ORF Triads differ from one another. Therefore, because the passages selected for progress monitoring are approximately equivalent to the passages selected for screening, the reliability estimates for the ORF Triad are approximately equivalent to the reliability estimates for ORF progress monitoring passages.

Reliability. The alternate-form reliability estimates listed in this document are from the *DIBELS Next Oral Reading Fluency Readability Study (Technical Report No. 7)* (see tables 24, 27, 30, 33, 36, and 39, pp. 65-80) (Powell-Smith, Good, & Atkins, 2010). These estimates are also available in the *Acadience Reading Technical Manual* (table 5.7, page 92) (Good et al., 2013). The alternate-form reliability reported is the median reliability based on all pair-wise grade-level passage correlations.

The estimates for test-retest and inter-rater reliability listed in this document were calculated during the *DIBELS Next Benchmark Goals Study* (Powell-Smith et al., 2012). In this study, inter-rater reliability was evaluated at beginning-of-year benchmark assessment. Students were given an ORF Triad and two individual raters assessed the student's responses. For test-retest reliability, the same ORF Triad was given to students approximately two weeks after middle-of-year benchmark assessment, and their scores were correlated. This information is also available in the *Acadience Reading Technical Manual* (see table 5.18, page 99) (Good et al., 2013).

Validity. The concurrent validity estimates listed in this document are from the *DIBELS Next Oral Reading Fluency Readability Study (Technical Report No. 7)* (see tables 24, 27, 30, 33, 36, and 39, pp. 65-80) and are represented by the correlations between all progress monitoring ORF passages and the Standard 4th Grade Reading Passage used in the NAEP 2002 Special Study of Oral Reading (Daane et al., 2005). The estimate reported is the median correlation based on all pair-wise grade-level passage correlations. Predictive validity estimates are from *DIBELS Next: Findings from the Benchmark Goals Study (Technical Report No. 11)* (see tables 31-36, pp. 105-128)

(Powell-Smith et al., 2012) and are represented by the correlation between the median ORF Triad passage score and the GRADE Total Test raw score (Williams, 2001). Predictive and concurrent validity is also available in the *Acadience Reading Technical Manual* (Table 6.13 and Table 6.14, pp. 110-111) (Good et al., 2013).

Growth Rate. Growth rate norms are represented by the mean of the estimated slope from a hierarchical linear regression (HLM) model predicting ORF Words Correct (WC) over time. The Growth Rate Criterion is separated into three different metrics: below typical, typical, and above typical growth. Below typical, typical, and above typical growth are represented by the 20th percentile rank, the 40th percentile rank, and the 60th percentile rank of the distribution of the estimated slope, respectively.

The reliability of the estimated individual growth rate, i.e., the reliability of the slope of improvement, was calculated from an HLM allowing both the slope and the intercept to vary across students. Additionally, two criteria were placed upon the data to ensure the integrity of the results. The purpose for setting these criteria was to gather a set of students for which calculating the slope was reasonable and defensible. The first criterion was to set the minimum number of data points for including students in the analysis. Previous work suggested that 14 data points was the minimum number of data points with the highest reliability (Good, 2009, and Good et al., 2010), and represented an adequate amount of data that would establish a trend that could be adequately modeled. Thus, we selected 14 as the minimum number of data points required for inclusion in the analysis. The second criterion was to select students for the analysis based on the root mean squared error (RMSE) of the HLM. The RMSE could be artificially inflated due to additional variability that is not explained by the student's scores such as environmental concerns (e.g., inadequate or uncomfortable facilities) or errors in data entry. Examples of student progress monitoring records for which there was suspicion of the integrity of the data are displayed in Figures 1 and 2. For first and second grades, students whose RMSE was less than 11 were included in the analysis. For third, fourth, fifth, and sixth grades, students whose RMSE was less than 10 were included in the analysis. With these criteria met, the analysis was conducted to calculate the reliability of the slope, (i.e., the estimated individual growth rate).

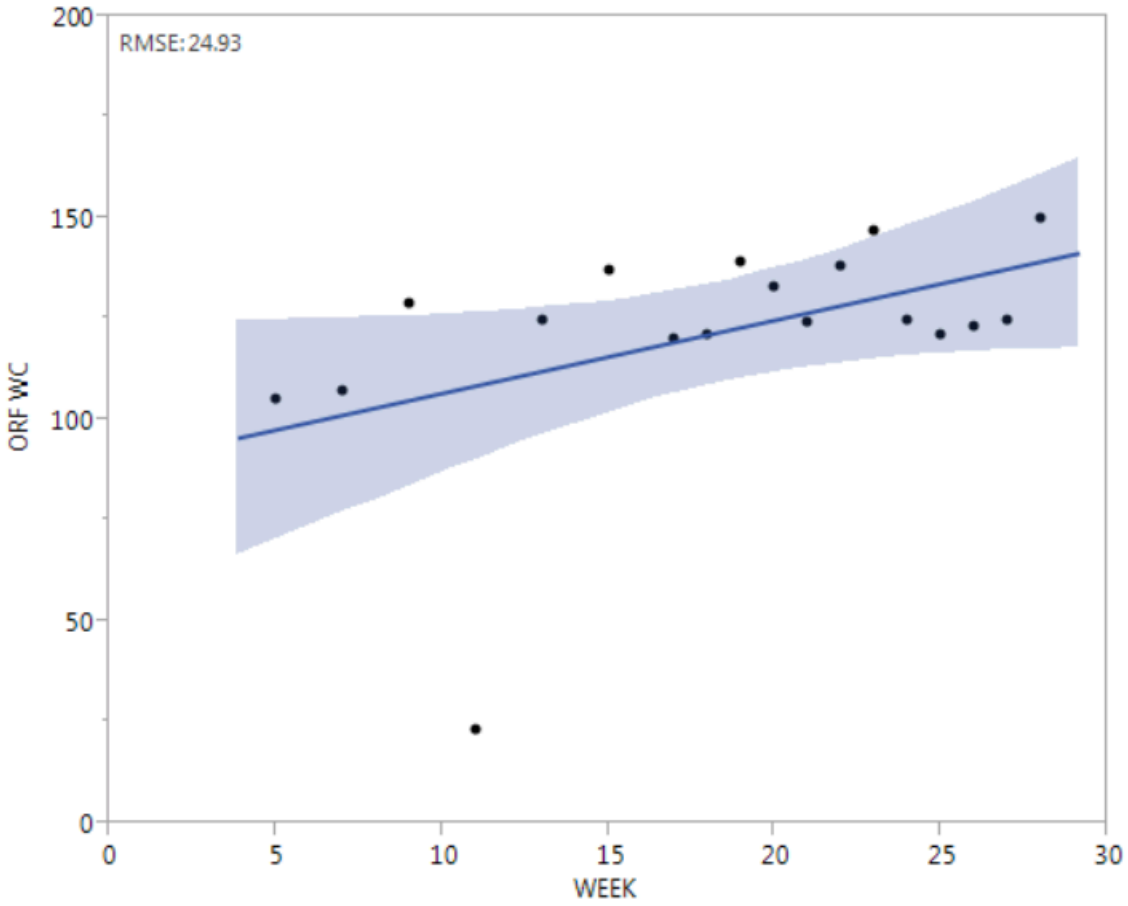

Figure 1. *A fourth-grade student progress monitoring record where the RMSE was unusually large. For Week 11, the student either did not perform the task as instructed or there was an error in data entry.*

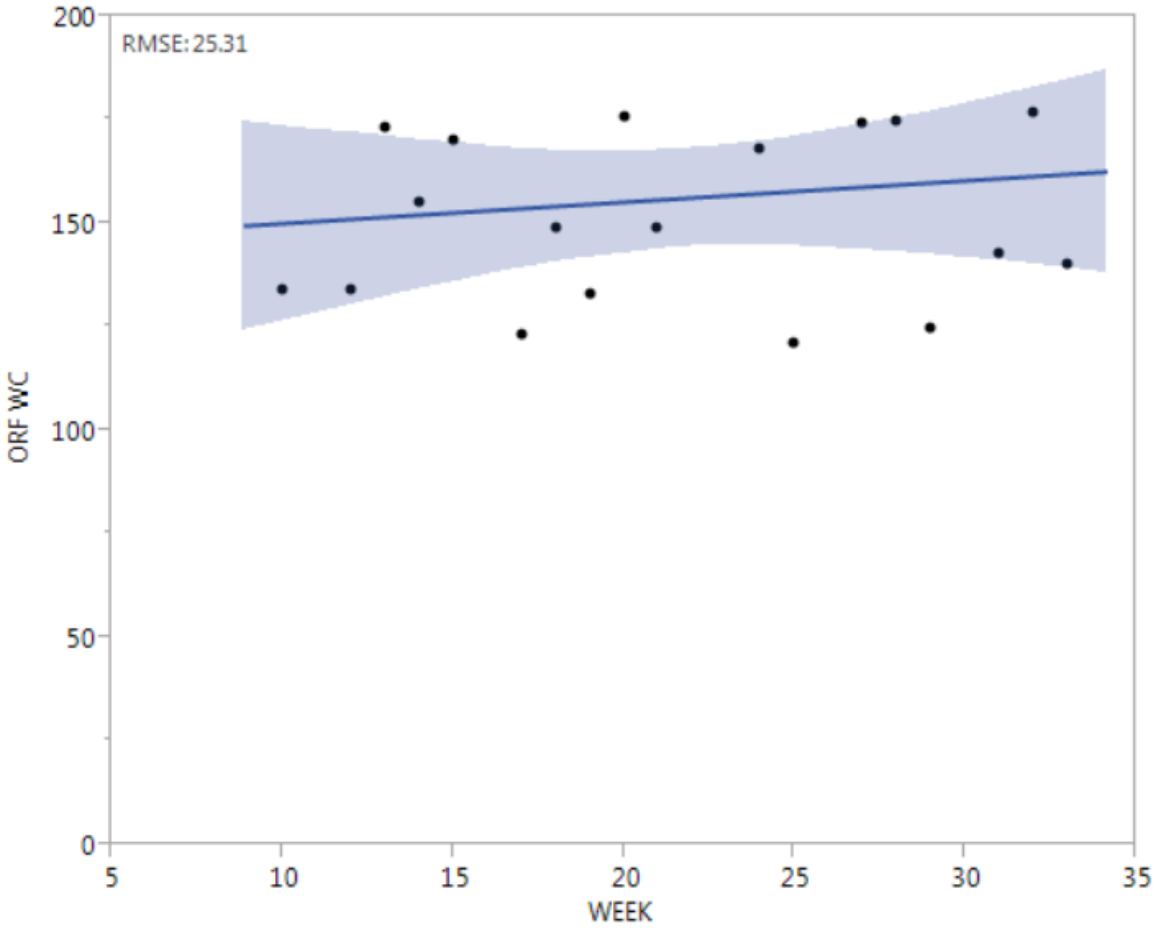

Figure 2. *A fourth-grade student progress monitoring record where the RMSE was unusually large. The student's responses across weeks vary widely enough (approximately 50 points) that using the slope as a guiding metric for growth is not defensible.*

Table 1. *Technical Adequacy for Acadience Reading Oral Reading Fluency Words Correct*

	Grade					
	1	**2**	**3**	**4**	**5**	**6**
Alternate-Form Reliability	.95	.91	.93	.90	.92	.84
Sample size (N)	23	25	22	23	23	24
Test-Retest Reliability	.95	.91	.93	.97	.97	--
Sample size (N)	28	21	27	21	23	--
Inter-Rater Reliability	--	.99	.99	.99	.99	.99
Sample size (N)	--	25	25	24	28	20
Predictive Validity	.64	.76	.67	.77	.69	.64
Sample size (N)	196	215	190	190	194	103
Concurrent Validity	.97	.91	.96	.89	.96	.83
Sample size (N)	23	25	22	23	23	24
Growth Rate Norms:	1.09	1.16	.61	.55	.45	.58
Sample size (N)	356	2051	843	1010	610	102
Growth Rate Criterion (Above Typical)	1.24	1.27	.68	.62	.51	.57
Growth Rate Criterion (Typical)	.94	1.05	.54	.48	.38	.45
Growth Rate Criterion (Below Typical)	.60	.78	.38	.30	.23	.30
Sample size (N)	356	2051	843	1010	610	102
Reliability of Estimated Individual Growth Rate	.82	.77	.55	.56	.50	.50
Sample size (N)	356	2051	843	1010	610	102
Reliability of Growth Rate	--	--	--	--	--	--
Sample size (N)	--	--	--	--	--	--
Single-Passage Standard Error of Measurement	10.33	11.29	11.12	10.5	10.39	10.96

Sample size (N)	23	25	22	23	23	24
Number of Passages	29	32	32	32	32	32

Note. Alternate-form reliability is the median reliability from all possible pair-wise correlations between 20 passages administered over four to seven days (two to five passages per day). Test-rest forms were given after an approximate two-week delay. Above typical, typical, and below typical growth rate criteria represent the 60th, 40th, and 20th percentile rank of the distribution of estimated individual slope, respectively.

References

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). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office. Available http://nces.ed.gov/nationsreportcard/pdf/studies/2006469.pdf. Accessed 6/22/2010.

Good, R. H. (2009, February). Evidentiary Requirements for Progress Monitoring Measures When Used for Response to Intervention. Paper presented at the DIBELS Summit, Albuquerque, NM.

Good, R. H., Kaminski, R. A., Dewey, E. N., Wallin, J., Powell-Smith, K. A., & Latimer, R. J. (2013). *DIBELS Next Technical Manual.* Eugene, OR: Dynamic Measurement Group. Retrieved from https://acadiencelearning.org/.

Good, R. H., Wheeler, C. E., Cummings, K. D., Baker, S. K., Fien, H., & Kame'enui, E. J. (2010, March). Rigorous RTI Decisions: Normative Growth Rates for Oral Reading Fluency. Paper presented at the National Association of School Psychologists Conference, Chicago, IL.

Powell-Smith, K. A., Good, R. H., & Atkins, T. (2010). *DIBELS Next Oral Reading Fluency Readability Study* (Technical Report No. 7). Eugene, OR: Dynamic Measurement Group. Retrieved from https://acadiencelearning.org/.

Powell-Smith, K. A., Good, R. H., Latimer, R. J., Dewey, E. N., Wallin, J., & Kaminski, R. A. (2012). *DIBELS Next: Findings from the Benchmark Goals Study (Technical Report No. 11).* Eugene, OR: Dynamic Measurement Group. Retrieved from https://acadiencelearning.org/.

Williams, K. T., (2001). Group Reading Assessment and Diagnostic Evaluation (GRADE). New York, NY: Pearson.