ACADIENCE LEARNING INC.

Acadience™ Reading K–6 Technical Adequacy Brief

Suggested Citation

Dewey, E. N., Powell-Smith, K. A., Good, R. H., Kaminski, R. A. (2015) *Acadience™ Reading K-6 Technical Adequacy Brief*. Eugene, OR: Acadience Learning.

Technical Adequacy Information Update for Acadience Reading K–6

Since the publication of Acadience Reading (also published as DIBELS Next) in 2010 and the publication of the *Acadience Reading Technical Manual* (Good et al., 2013), a number of new studies have been conducted examining the technical adequacy of Acadience Reading. The results of these studies have been published in various forms (websites, papers, etc.) and have not yet been integrated into the *Technical Manual* (Good et al., 2013). Until these data can be integrated, this document will provide a quick reference to the most up-to-date information on the technical adequacy (e.g., reliability and validity) of Acadience Reading. For a timeline showing which Acadience Reading measures are administered at each benchmark assessment period for all grades, see *Appendix 1* at the end of this document.

What's new?

Alternate-Form Reliability. Updated alternate-form reliability is available for First Sound Fluency (FSF), Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency Correct Letter Sounds (NWF CLS) and Whole Words Read (NWF WWR), and Maze. These new estimates are reported in place of the previous estimates in a summary table that includes all reliability information for Acadience Reading.

Reliability of the Slope. New estimates for the reliability of the individual growth rate (slope) is available for all Acadience Reading measures in all grades.

Descriptions of the Samples

Technical adequacy information for Acadience Reading was calculated from data sets exported from the Acadience Data Management and mCLASS® data reporting services. The data were collected and entered by school personnel. Every census region in the United States was represented. There are five primary samples.

Sample 1. Progress monitoring data from the school year 2011-2012 were exported from the mCLASS data reporting service. There were approximately 3,000 students in third through

sixth grade. The sample was approximately 61% White, 15% African American, and 13% Hispanic.

Sample 2. Progress monitoring and benchmark data from the 2013-2014 school year were exported from the mCLASS data reporting service. There were approximately 21,157 students in kindergarten and first grade from 2,196 schools within 634 districts. The sample was approximately 45% White, 20% African American, and 27% Hispanic.

Sample 3. Progress monitoring data from the school year 2013-2014 were exported from Acadience Data Management. There were approximately 130,000 students in first through sixth grades from 1,121 schools within 388 school districts. Demographic information for this sample is not available.

Sample 4. Benchmark administration, validity, and reliability data were collected during the 2009-2010 school year. There were approximately 3,816 students in kindergarten through sixth grade from thirteen schools within five school districts. Demographic information for this sample is not available. For more information about this sample, please refer to the *DIBELS Next Benchmark Goals Study*[1] (Powell-Smith et al., 2012), available for download from https://acadiencelearning.org/pubs.html.

Sample 5. Benchmark administration data from the 2012-2013 school year were exported from the mCLASS data reporting service. There were approximately 1,151,329 students from kindergarten through third-grade. The sample was approximately 25% White, 17% Black, 21% Hispanic, and 32% Unknown.

Analyses

All data were evaluated for validity, and invalid scores were removed. Invalid scores were classified as scores that are not possible, such as exceeding the maximum possible score, or pairs of invalid scores such as the sum of Oral Reading Fluency Words Correct (ORF WC) and

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

ORF Errors scores exceeding the maximum possible. For the reliability of the slope analysis, the student-level Root Mean Square Error (RMSE) was evaluated as an indicator of invalid data. For example, 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. Thus, students were included in the analysis if their RMSE was below 13 for FSF, 10 for PSF, 13 for NWF CLS, and 6 for NWF WWR. For ORF, students in first and second grades whose RMSE was less than 11 and students in third, fourth, fifth, and sixth grades whose RMSE was less than 10 were included in the analysis. Further, progress monitoring records that contained zero scores were removed for several reasons: 1) for ORF, the possibility that a student read zero words (including the connector words like "of", "and", and "the") is extremely unlikely, and it is more likely that the student did not attempt the task, and the assessor recorded a zero instead of a missing value; 2) a score of zero may indicate that the student did not understand the task and thus did not attempt it; and 3) a score of zero would likely be a leverage point and could alter the slope dramatically, thereby compromising the results of the analysis. Additionally, some students recorded multiple zeroes in combination with very low scores, in which case, progress monitoring on below-grade-level materials is recommended.

Reliability of the Slope of Improvement

The reliability of the estimated individual growth rate (i.e., the reliability of the slope of improvement) was calculated using Hierarchical Linear Modeling (Bryk & Raudenbush, 1992) allowing both the slope and the intercept to vary across students. Previous work suggested that students with at least 14 recorded data points over 36 weeks would provide a stable and reasonable estimate (Good, 2009; Good, et al., 2010). For all measures except Maze, a minimum of 14 data points was used to calculate slope (range = 14 to 40). For Maze, a minimum of 6 data points was used to calculate slope (range = 6 to 31). Reliability of the slope is reported in Table 1.

Reliability for Benchmark Administration

For alternate-form and test-retest reliability, all measures were given within a two-week time period and their scores were correlated. For evaluating scoring reliability, or rater-agreement, randomly selected students were administered the measures and were scored simultaneously by two assessors. Reliability for the alternate-form, test-retest, and inter-rater reliability is reported in Table 2.

Criterion-Related Validity

Concurrent and predictive criterion-related validity of Acadience Reading is presented as the correlation between the Reading Composite Score and the Group Reading Assessment and Diagnostic Evaluation (GRADE) (Williams, 2001) administered at the end of the 2009-2010 school year. For more information about criterion-related validity of Acadience Reading, please see the *Acadience Reading Technical Manual* (Good et al., 2013), available for download from https://acadiencelearning.org/pubs.html. Criterion-related validity is reported in Table 3.

Additional Technical Adequacy Information

The standard error of measurement is reported in Table 4. Sensitivity, specificity, and the area under the ROC analysis curve (AUC) are reported in Table 5. Internal consistency reliability for kindergarten through third-grade for the Reading Composite Score is reported in Table 6.

Table 1. *Reliability for the Slope of Improvement for Acadience Reading*

	Grade						
Acadience Reading Measure	K	1	2	3	4	5	6
First Sound Fluency	**.90**	--	--	--	--	--	--
Sample size (N)	2298	--	--	--	--	--	--
Phoneme Segmentation Fluency	**.86**	**.88**	--	--	--	--	--
Sample size (N)	1752	509	--	--	--	--	--
NWF Correct Letter Sounds	**.86**	**.87**	**.83**	--	--	--	--
Sample size (N)	779	15214	1555	--	--	--	--
NWF Whole Words Read	**.90**	**.88**	**.85**	--	--	--	--
Sample size (N)	292	14851	1550	--	--	--	--
ORF Words Correct	--	**.82**	**.77**	**.55**	**.56**	**.50**	**.50**
Sample size (N)	--	356	2051	843	1010	610	102
Maze	--	--	--	**.62**	**.61**	**.42**	**.35**
Sample size (N)	--	--	--	1562	471	396	570

Note. Based on Samples 1, 2, and 3 data. A '--' indicates the measure is not given at the specified time of year to the grade level.

Table 2. *Inter-Rater, Alternate-Form, and Test-Retest Reliability Estimates for Acadience Reading*

	Grade						
Acadience Reading Measure	K	1	2	3	4	5	6
First Sound Fluency							
Inter-Rater	.94	--	--	--	--	--	--
Single-Form Alternate-Form	.85	--	--	--	--	--	--
Three-Form Alternate-Form	.95	--	--	--	--	--	--
Single-Form Test-Retest	NA	--	--	--	--	--	--
Three-Form Test-Retest	NA	--	--	--	--	--	--
Phoneme Segmentation Fluency							
Inter-Rater	.96	.95	--	--	--	--	--
Single-Form Alternate-Form	.84	.83	--	--	--	--	--
Three-Form Alternate-Form	.94	.93	--	--	--	--	--
Single-Form Test-Retest	NA	NA	--	--	--	--	--
Three-Form Test-Retest	NA	NA	--	--	--	--	--
NWF Correct Letter Sounds							
Inter-Rater	.99	.99	NA	--	--	--	--
Single-Form Alternate-Form	.84	.85	.82	--	--	--	--
Three-Form Alternate-Form	.94	.94	.93	--	--	--	--
Single-Form Test-Retest	NA	.76	NA	--	--	--	--
Three-Form Test-Retest	NA	.90	NA	--	--	--	--
NWF Whole Words Read							
Inter-Rater	.99	.99	NA	--	--	--	--

Single-Form Alternate-Form	.92	.90	.83	--	--	--	--
Three-Form Alternate-Form	.97	.96	.93	--	--	--	--
Single-Form Test-Retest	NA	.70	NA	--	--	--	--
Three-Form Test-Retest	NA	.88	NA	--	--	--	--
ORF Words Correct							
Inter-Rater	--	NA	.99	.99	.99	.99	.99
Single-Form Alternate-Form	--	.95	.91	.93	.90	.92	.84
Three-Form Alternate-Form	--	.98	.96	.97	.95	.96	.90
Single-Form Test-Retest	--	.95	.91	.93	.97	.97	NA
Three-Form Test-Retest	--	.98	.97	.98	.99	.99	NA
Maze							
Inter-Rater	--	--	--	.99	.98	.99	.99
Single-Form Alternate-Form	--	--	--	.76	.79	.77	.79
Three-Form Alternate-Form	--	--	--	.90	.92	.91	.92
Single-Form Test-Retest	--	--	--	NA	NA	NA	NA
Three-Form Test-Retest	--	--	--	NA	NA	NA	NA
Reading Composite Score							
Inter-Rater	.97	.99	.98	NA	NA	NA	NA
Single-Form Alternate-Form	.66	.95	.92	.97	.95	.91	.91
Three-Form Alternate-Form	.85	.98	.97	.99	.98	.97	.97
Single-Form Test-Retest	NA	.94	.81	NA	NA	NA	NA
Three-Form Test-Retest	NA	.98	.93	NA	NA	NA	NA

Note. Based on data from Samples 1, 2, 3, and 4. NA = Not available. A '--' indicates the measure is not given to the grade level. Alternate-form reliability for kindergarten measures FSF, PSF, NWF CLS, first-grade PSF, second-grade NWF CLS and NWF WWR, and Maze in all grades is

calculated from progress monitoring data that was administered over approximately 36 weeks; the correlation represents the median reliability from all possible pair-wise assessments that were given in consecutive weeks. Alternate-form reliability for ORF 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). Alternate-form reliability forms for kindergarten LNF, NWF WWR, and first-grade NWF, and all test-retest forms were given after an approximate two-week delay after middle-of-year benchmark assessment. Test-retest reliability unavailable for Maze, all kindergarten and sixth-grade measures.

Table 3. *Predictive and Concurrent Criterion-Related Validity for Acadience Reading*

	Grade						
Acadience Reading Measure	K	1	2	3	4	5	6
First Sound Fluency							
Predictive	.52	--	--	--	--	--	--
Concurrent	NA	--	--	--	--	--	--
Phoneme Segmentation Fluency							
Predictive	.34	.33	--	--	--	--	--
Concurrent	.24	NA	--	--	--	--	--
NWF Correct Letter Sounds							
Predictive	.47	.51	.51	--	--	--	--
Concurrent	.40	.56	NA	--	--	--	--
NWF Whole Words Read							
Predictive	.19	.52	.51	--	--	--	--
Concurrent	.35	.56	NA	--	--	--	--
ORF Words Correct							
Predictive	--	.64	.76	.67	.77	.65	.59
Concurrent	--	.75	.73	.66	.74	.65	.61
Maze							
Predictive	--	--	--	.65	.67	.56	.60
Concurrent	--	--	--	.67	.68	.66	.64
Reading Composite Score							
Predictive	.48	.71	.80	.78	.80	.76	.68

Concurrent	.40	.77	.75	.75	.80	.77	.73

Note. Based on Sample 4 data. Total sample size = 1306. NA = Not available. A '--' indicates the measure is not given to the grade level. Acadience Reading validity based on correlations with the Group Reading Assessment and Diagnostic Evaluation (GRADE) Total Test Raw Score which was administered at end-of-year benchmark administration. Concurrent validity is not currently available for FSF, first-grade PSF, and second-grade NWF, because those measures are not administered at the end of the year.

Table 4. *Standard Error of Measurement of Acadience Reading Measures*

	Grade						
Acadience Reading Measure	K	1	2	3	4	5	6
First Sound Fluency	4.68	--	--	--	--	--	--
Phoneme Segmentation Fluency	7.57	5.16	--	--	--	--	--
NWF Correct Letter Sounds	5.82	12.59	5.70	--	--	--	--
NWF Whole Words Read	1.15	4.27	2.36	--	--	--	--
ORF Words Correct	--	5.56	8.00	7.00	8.53	7.66	7.00
Maze	--	--	--	3.91	4.00	4.68	2.95
Reading Composite Score	28.46	22.35	28.23	20.69	25.17	31.57	15.19

Note. SEM calculated using alternate-form reliability estimates from data in Samples 2 and 4. A '--' indicates the measure is not given to the grade level.

Table 5. *Sensitivity, Specificity, and Area Under the Curve (AUC) for Acadience Reading Measures*

	Grade						
Acadience Reading Measure	K	1	2	3	4	5	6
First Sound Fluency							
Sensitivity	.34	--	--	--	--	--	--
Specificity	.88	--	--	--	--	--	--
AUC	.72	--	--	--	--	--	--
Phoneme Segmentation Fluency							
Sensitivity	.09	.24	--	--	--	--	--
Specificity	.89	.75	--	--	--	--	--
AUC	.56	.65	--	--	--	--	--
NWF Correct Letter Sounds							
Sensitivity	.24	.50	.47	--	--	--	--
Specificity	.75	.83	.78	--	--	--	--
AUC	.69	.80	.86	--	--	--	--
NWF Whole Words Read							
Sensitivity	NG	.50	.53	--	--	--	--
Specificity	NG	.81	.73	--	--	--	--
AUC	NG	.78	.82	--	--	--	--
ORF Words Correct							
Sensitivity	--	.67	.75	.61	.69	.61	.38
Specificity	--	.92	.89	.87	.79	.83	.90
AUC	--	.92	.89	.87	.87	.83	.85

Maze							
Sensitivity	--	--	--	.65	.69	.48	.25
Specificity	--	--	--	.88	.78	.78	.90
AUC	--	--	--	.86	.80	.82	.84
Reading Composite Score							
Sensitivity	.10	.70	.75	.74	.77	.68	.50
Specificity	.85	.88	.89	.90	.84	.88	.93
AUC	.66	.90	.88	.90	.88	.88	.92

Note. Based on Sample 4 data. A '--' indicates the measure is not given to the grade level. NG = no benchmark goal for that grade-level.

Table 6. *Internal Consistency Reliability for Acadience Reading Composite Score*

Reading Composite Score by Time of Year	Grade			
	K	1	2	3
Beginning of Year	NA	.76	.89	.85
Middle of Year	.83	.90	.83	.84

Note. Based on Sample 5 data. NA = Not Available. Sample size approximately N = 1,150,000.

Appendix 1. Timeline of Administration for Acadience Reading Measures

Acadience Reading Measure and Time of Year	Grade						
	K	1	2	3	4	5	6
FSF							
Beginning of Year	**X**	--	--	--	--	--	--
Middle of Year	**X**	--	--	--	--	--	--
End of Year	--	--	--	--	--	--	--
PSF							
Beginning of Year	--	**X**	--	--	--	--	--
Middle of Year	**X**	--	--	--	--	--	--
End of Year	**X**	--	--	--	--	--	--
NWF CLS							
Beginning of Year	--	**X**	**X**	--	--	--	--
Middle of Year	**X**	**X**	--	--	--	--	--
End of Year	**X**	**X**	--	--	--	--	--
NWF WWR							
Beginning of Year	--	**X**	**X**	--	--	--	--
Middle of Year	**X**	**X**	--	--	--	--	--
End of Year	**X**	**X**	--	--	--	--	--
ORF WC							
Beginning of Year	--	--	**X**	**X**	**X**	**X**	**X**
Middle of Year	--	**X**	**X**	**X**	**X**	**X**	**X**
End of Year	--	**X**	**X**	**X**	**X**	**X**	**X**

Retell							
Beginning of Year	--	--	**X**	**X**	**X**	**X**	**X**
Middle of Year	--	**X**	**X**	**X**	**X**	**X**	**X**
End of Year	--	**X**	**X**	**X**	**X**	**X**	**X**
ORF Accuracy							
Beginning of Year	--	--	**X**	**X**	**X**	**X**	**X**
Middle of Year	--	**X**	**X**	**X**	**X**	**X**	**X**
End of Year	--	**X**	**X**	**X**	**X**	**X**	**X**
Maze							
Beginning of Year	--	--	--	**X**	**X**	**X**	**X**
Middle of Year	--	--	--	**X**	**X**	**X**	**X**
End of Year	--	--	--	**X**	**X**	**X**	**X**

Note. A '--' indicates the measure is not given at the specified time of year to the grade level. FSF = First Sound Fluency, PSF = Phoneme Segmentation Fluency, NWF CLS = Nonsense Word Fluency Correct Letter Sounds, NWF WWR = Nonsense Word Fluency Whole Words Read, ORF WC = Oral Reading Fluency Words Correct, ORF Accuracy = Oral Reading Fluency Accuracy.

References

Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical Linear Models in Social and Behavioral Research: Applications and Data Analysis Methods* (First Edition). Newbury Park, CA: Sage Publications.

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 from http://nces.ed.gov/nationsreportcard/pdf/studies/2006469.pdf. Accessed 6/22/2010.

Good, R. H., Kaminski, R. A., Dewey, E. N., Wallin, J., Powell-Smith, K. A., & Latimer, R. J. (2013). *Acadience Reading Technical Manual.* Eugene, OR: Acadience Learning. Available: 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.

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.

Gushta, M., & Wang, Y. (2013) Exploring *DIBELS Next* Composite Score Performance by Student Subgroups. Presented at the *DIBELS* Literacy Summit on March 6, 2013.

Hopkins, W. G. (2002). A scale of magnitudes for the effect statistics. In A *review of statistics*. Retrieved July 30, 2010 from http://www.sportsci.org/resource/stats/effectmag.html

Hox, J.J. & van de Schoot, R. (2013). Robust methods for multilevel analysis. Pp. 387-402 in M.A. Scott, J.S. Simonoff & B.D. Marx (Eds.) *The Sage Handbook of Multilevel Modeling*. Los Angeles: Sage. Retrieved May 27, 2014 from http://joophox.net/publist/Chap14ProofsX.pdf

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

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