DIBELS® 8TH EDITION

UNIVERSITY OF OREGON | College of Education

Dynamic Indicators of Basic Early Literacy Skills

8th Edition

Improved DIBELS® 8th Edition Benchmark Goals Technical Evidence: A DIBELS 8th Edition Technical Manual Supplement

Contributing Authors

Gina Biancarosa, Ed. D.
Associate Professor and Ann Swindells Chair in Education
Department of Educational Methodology, Policy, and Leadership
College of Education, University of Oregon

Patrick C. Kennedy, Ph.D.
Research Associate and Director of Data Management and Analysis Group
Center on Teaching & Learning
College of Education, University of Oregon

Christopher Ives
Graduate Research Assistant Fellow
Center on Teaching & Learning
College of Education, University of Oregon

Contributing Editors

Maureen Warman, M.S.
Senior Research Assistant II and DIBELS Data System Manager
Center on Teaching & Learning
College of Education, University of Oregon

Table of Contents

Introduction

This technical manual supplement describes updates to DIBELS 8th Edition (DIBELS 8) benchmark goals for the 2020-2021 school year and beyond. As described in the 8th Edition of DIBELS: Technical Manual, the benchmark goals published with the release of DIBELS 8 were established using (a) receiver operating characteristic (ROC) curve analyses from the 2018-2019 validity study, (b) criteria set by the National Center on Intensive Intervention (NCII), and (c) percentile rank estimates from the 2017-2018 and 2018-2019 research samples. The updates described here incorporate data collected in 2019-2020 on nearly one million additional students from across the country whose scores were entered into either the DIBELS Data System (DDS) or our mobile partner Amplify's mCLASS system.

These updates improve the classification concordance (i.e., agreement) between subtest and composite scores and calibrate the goals to align with new national percentile rank estimates derived from the most recent operational sample. These updates are based on the original ROC analyses, so overall classification accuracy of the benchmark goals remains unchanged, but the changes improve the precision with which we identify those students who are most at risk for not meeting grade-level goals.

In this document, we provide more information on the motivation, procedures, and specific changes underlying the benchmark goal updates, as well as tables of classification accuracy for the updated goals compared to previous goals.

Chapter 1: Motivations and Procedures

Need for Updates

Changes to the DIBELS 8 benchmark goals were motivated by two primary factors: (a) the concordance (i.e., agreement) between subtest and composite risk classifications for individual students, and (b) percentile ranks updated with the latest operational data from DDS and Amplify.

With the widespread implementation of DIBELS 8 in the 2019-2020 school year, some users reported instances in which students' composite risk classifications were discrepant from their classifications on DIBELS 8 subtests. For example, we discovered that, in some grades, it was possible for students' scores to be at-benchmark (i.e., green) on each subtest, yet receive a composite score corresponding to the strategic support (i.e., below benchmark/yellow) category. This lack of concordance was a consequence of setting cut-points for each subtest score and composite score separately. That is, benchmark goals were originally selected to optimize each individual cut-point's sensitivity and specificity, without reference to the alignment between subtests and composite scores. This lack of concordance unnecessarily complicated decision-making for users, and consequently, we updated benchmark goals to improve subtest and composite concordance without compromising the accuracy of identifying students who are at-risk or not at-risk.

Our review of the updated percentile ranks revealed that although our research sample typically yielded percentile ranks that were similar to those in the 2019-2020 operational dataset, there were exceptions. For example, the 50th percentile in the research sample may have turned out to be the 60th percentile in the operational sample. As a result, we further refined benchmark goals by calibrating them to the updated percentile ranks to ensure that the percentage of students designated at risk approached what schools implementing multi-tiered systems of support (MTSS) expect and can reasonably support (i.e., about 20 percent of students in need of intensive intervention and 40 percent in need of strategic support). Again, the strong relationship between DIBELS 8 and the criterion measures allowed us to make these refinements without compromising classification accuracy.

Additional Considerations During New Cut-Score Selection

Our methods for benchmark goals refinement utilized the original ROC analyses described in the Technical Manual. Thus, the overall classification accuracy (i.e., area under the curve, or AUC) of they benchmark goals is unchanged. Our process built upon the framework described in the Technical Manual by invoking deeper consideration of cut-score concordance, updated percentile ranks, and the developmental progression across grades, all while continuing to follow the industry-standard NCII criteria for screening measures.

To satisfy each of these considerations, we employed an iterative evaluation process. First, we identified the range of scores that satisfied the NCII criteria for convincing evidence of classification accuracy, which include (a) an AUC confidence interval that is .80 or higher, (b) sensitivity of .70 or higher, and (c) specificity of .80 or higher (as well as an appropriate base rate in the research sample).[1] To illustrate this, Figure 1 depicts in blue the 17 scores for the Grade 2 end of year composite score that meet NCII criteria for convincing evidence. In general, the situation depicted in Figure 1 was similar for composite scores in most grades and times of year with multiple scores from which to choose the cut-point.
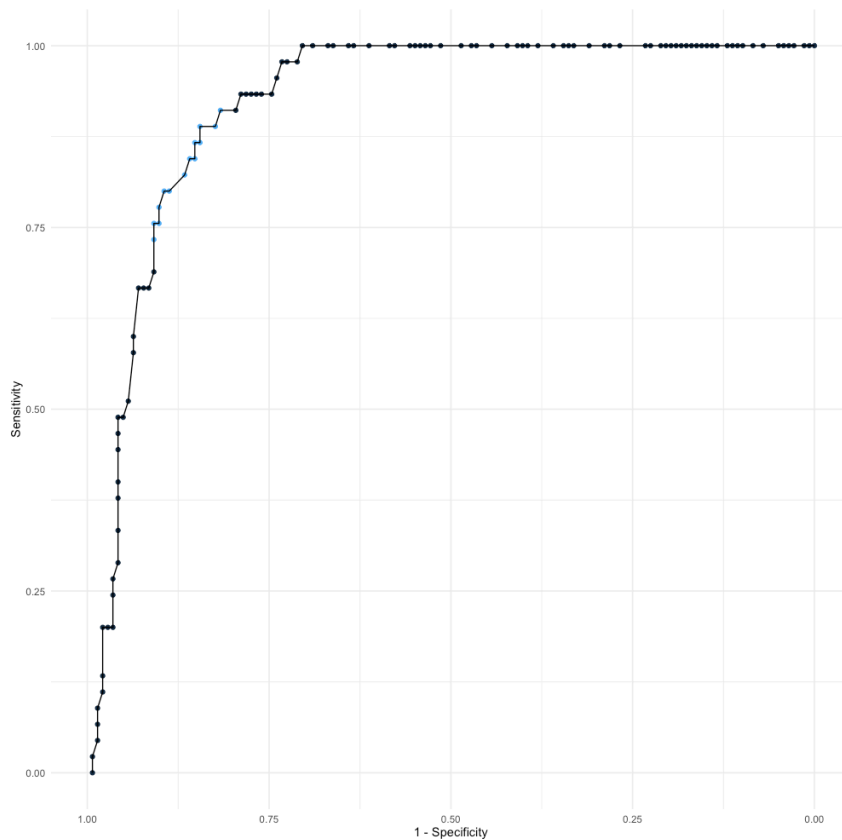

Figure 1. ROC curve demonstrating a range of scores meeting NCII criteria for the Grade 2 end of year composite score, where blue dots represent NCII convincing evidence.

[1] In cases where convincing evidence was not quite achieved, we identified the range of scores meeting the criteria for partially convincing evidence, etc.

The range of potential benchmark goals were then carefully examined to maximize each area of consideration: the associated percentile rank, concordance between a subtest and the composite score, and appropriate developmental progression within and across grades. Where percentile ranks were found to be too high, we adjusted cut-points down, which resulted in a smaller percentage of students being classified in the category below the cut and improving specificity, which represents the percentage of students above the cut who can be expected to meet or exceed the criterion, while simultaneously reducing the number of false positives (i.e., the number of students identified as at risk who are not at risk). This privileging of specificity without sacrificing too much sensitivity is also more in keeping with the spirit of NCII evidence criteria, which holds that to support effective resource allocation for students in need of intensive and strategic support, cut-points should have a higher threshold for specificity than for sensitivity.

As in Figure 1, many cut-points for classifying students as at risk and in need of intensive intervention can demonstrate adequate screening accuracy based on NCII criteria. Nonetheless, in this figure and based on percentile ranks, the lowest cut-point (i.e., 415) would identify 17% of students nationally as being at risk and the highest (i.e., 432) would identify 30%. For this score, we ultimately chose 420 as the cut-point, which would identify 21% of students nationally. In general, the updated DIBELS 8 benchmark goals seek to align the strategic and intensive range goals to the 40th and 20th percentiles, respectively.

The Unique Case of ORF Accuracy

The only exception to the procedures just described was how we set updated benchmark goals for ORF Accuracy scores. Previous benchmark goals were set primarily based on NCII criteria; however, the adherence to the criteria in this case resulted in extremely high cut-points in Grades 2 and beyond that do not reflect instructional expectations. In addition, ORF accuracy is not a strong predictor of risk beyond first grade, as reflected in the formulas for composite score calculations. Nevertheless, ORF accuracy provides rich formative data that can be used in designing instructional interventions.

As a result, we changed ORF accuracy cut-points beyond the beginning of first grade to adhere to instructional norms to improve their instructional utility. While individual curriculum and studies differ in where they draw these cut-points (cf., Betts, 1946; Halladay, 2012; Treptow, Burns, & McComas, 2007), it is generally agreed that 90% or less accuracy likely indicates a student will be challenged to comprehend a text at that level, while greater than 95% indicates a student should experience little to no challenge in comprehending a text at that level. Importantly, this research varies in how they define the cut-points and how they judge "frustration." Treptow et al. (2007) found that students reading instructional level texts, which they defined as 93-97% accuracy, showed greater on-task behavior than students reading at their frustration and independent levels.

For the purposes of DIBELS 8 ORF, from the middle of first grade through the end of eighth grade, a student with 96% accuracy or above is classified as at minimal risk for struggling to comprehend grade-level texts. A student scoring between 91% and 95% accuracy inclusively is classified as at some risk and in need of strategic support for comprehending grade level texts, while a student scoring 90% accuracy or below is at risk and requires intensive support for comprehending grade level texts.

Chapter 2: Technical Adequacy of New DIBELS 8th Edition Benchmark Goals

In the following pages, we present tables comparing the sensitivity and specificity for the updated goals to those values for the previous goals. As noted, we used the original ROC analyses, so the overall classification accuracy of benchmark goals (i.e., the area under the curve, or AUC) is unchanged, but adjusting benchmark goal cut-points improved the precision with which we identified students most at risk for not meeting grade-level goals by increasing specificity and lowering false positive rates. In some cases, but not all, the updated goals may have slightly reduced sensitivity. And in a very few cases, in the upper grades, sensitivity may be more seriously reduced. We judged this acceptable given that reading comprehension as assessed by measures like the Iowa Test of Reading involves more complex thinking than in lower grades, limiting (although not eliminating) the sensitivity of these measures to the wider range of sources of reading comprehension difficulty in the upper grades. Nonetheless, we only made such updates if we could vastly improve specificity.

In many cases, sensitivity and specificity are identical because goals did not change. Where goals changed, users can see how sensitivity and specificity were affected. What the tables make apparent is that updated goals generally improved specificity, although in some cases it was sensitivity that was improved. Improving specificity results in fewer false positives, or students designated as at risk who are not, in fact, at risk. In contrast, improving sensitivity results in greater true positives, or students designated as at risk who are, in fact, at risk.

We begin with tables for the composite score cut-points because these are the best overall indicator of risk regarding achieving end of year learning goals. Based on cut-point changes, the risk classification on composite scores is now much more likely to be consistent (or concur) with risk classifications on subtests. Thus, performance on the subtests can be viewed as a more diagnostic view of how a student was classified and relative areas of strength and weakness. Following the composite scores, we report on the subtests by grade level: from kindergarten through eighth grade.

Table 2.1 Receiver Operating Curve Results for DIBELS 8 Composite Predicting End-of-year Criterion Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
K	20th	1	306	0.71	0.88	0.71	0.88
		2	309	0.78	0.87	0.73	0.90
		3	321	0.80	0.93	0.80	0.93
	40th	1	306	0.65	0.83	0.65	0.83
		2	309	0.80	0.85	0.77	0.86
		3	321	0.88	0.84	0.78	0.93
1	20th	1	112	0.81	0.79	0.65	0.86
		2	135	0.77	0.79	0.80	0.79
		3	128	0.77	0.88	0.80	0.86
	40th	1	112	0.78	0.69	0.72	0.76
		2	135	0.83	0.68	0.84	0.63
		3	128	0.80	0.78	0.80	0.78
2	20th	1	127	0.82	0.86	0.82	0.89
		2	149	0.82	0.89	0.81	0.89
		3	187	0.76	0.90	0.80	0.89
	40th	1	127	0.78	0.77	0.78	0.79
		2	149	0.78	0.85	0.78	0.81
		3	187	0.81	0.85	0.81	0.85
3	20th	1	109	0.63	0.78	0.63	0.78
		2	165	0.65	0.78	0.65	0.79
		3	171	0.70	0.80	0.70	0.82
	40th	1	109	0.71	0.69	0.71	0.69
		2	165	0.80	0.77	0.76	0.80
		3	171	0.80	0.77	0.80	0.79
4	20th	1	129	0.71	0.80	0.71	0.80
		2	182	0.85	0.88	0.85	0.88
		3	180	0.73	0.86	0.73	0.86
	40th	1	129	0.74	0.64	0.67	0.79
		2	182	0.80	0.75	0.70	0.90
		3	180	0.91	0.61	0.68	0.79

Table 2.1 Receiver Operating Curve Results for DIBELS 8 Composite Predicting End-of-year Criterion Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
5	20th	1	95	0.71	0.79	0.71	0.79
		2	133	0.76	0.86	0.76	0.86
		3	109	0.71	0.83	0.71	0.83
	40th	1	95	0.82	0.60	0.76	0.68
		2	133	0.86	0.74	0.80	0.82
		3	109	0.82	0.74	0.79	0.77
6	20th	1	44	0.83	0.68	0.67	0.82
		2	102	0.78	0.82	0.67	0.89
		3	96	0.63	0.74	0.63	0.89
	40th	1	44	0.84	0.52	0.63	0.64
		2	102	0.93	0.67	0.67	0.84
		3	96	0.83	0.67	0.46	0.74
7	20th	1	36	0.90	0.96	0.90	0.96
		2	92	0.75	0.85	0.75	0.88
		3	94	0.77	0.86	0.69	0.90
	40th	1	36	0.78	0.83	0.78	0.83
		2	92	0.96	0.80	0.96	0.80
		3	94	0.79	0.83	0.71	0.84
8	20th	1	45	0.81	0.83	0.76	0.83
		2	49	0.70	0.81	0.78	0.77
		3	52	0.71	0.79	0.75	0.61
	40th	1	45	0.78	0.77	0.78	0.77
		2	49	0.77	0.71	0.80	0.71
		3	52	0.68	0.60	0.68	0.60

Note. Criteria were percentile ranks. Criterion measure in kindergarten was DIBELS Next Composite score. Criterion measure in Grades 1-8 was Iowa Assessment Total Reading score. 1 = Beginning of year. 2 = Middle of year. 3 = End of year.

Table 2.2 Receiver Operating Curve Results for DIBELS 8 Kindergarten Subtests Predicting End-of-year DIBELS Next Composite Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
LNF	20th	1	306	0.75	0.85	0.73	0.85
		2	314	0.75	0.78	0.71	0.80
		3	321	0.82	0.79	0.78	0.82
	40th	1	306	0.73	0.76	0.71	0.79
		2	314	0.78	0.74	0.77	0.76
		3	321	0.85	0.80	0.84	0.82
PSF	20th	1	306	0.69	0.73	0.39	0.93
		2	309	0.75	0.85	0.71	0.85
		3	321	0.76	0.82	0.75	0.84
	40th	1	306	0.83	0.58	0.45	0.87
		2	309	0.72	0.78	0.64	0.84
		3	321	0.74	0.68	0.61	0.75
NWF-CLS	20th	1	306	0.67	0.78	0.65	0.81
		2	309	0.75	0.84	0.71	0.85
		3	321	0.78	0.89	0.75	0.91
	40th	1	306	0.80	0.71	0.75	0.78
		2	309	0.82	0.77	0.73	0.81
		3	321	0.78	0.85	0.73	0.87
NWF-WRC	20th	1	306	>.99	0.40	NA	NA
		2	309	0.86	0.77	0.65	0.86
		3	321	0.75	0.86	0.71	0.89
	40th	1	NA	NA	NA	0.91	0.56
		2	309	0.79	0.78	0.67	0.85
		3	321	0.80	0.82	0.70	0.89
WRF	20th	1	306	0.86	0.43	NA	NA
		2	309	0.75	0.86	0.71	0.90
		3	321	0.78	0.87	0.73	0.87
	40th	1	NA	NA	NA	0.76	0.53
		2	309	0.77	0.79	0.69	0.86
		3	321	0.79	0.82	0.72	0.85

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year. NA = Not applicable (i.e., no cut score for this period).

Table 2.3 Receiver Operating Curve Results for DIBELS 8 First Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
LNF	20th	1	123	0.68	0.62	0.50	0.82
		2	137	0.71	0.64	0.63	0.74
		3	134	0.72	0.68	0.53	0.77
	40th	1	123	0.80	0.41	0.57	0.70
		2	137	0.70	0.56	0.63	0.70
		3	134	0.67	0.63	0.58	0.75
PSF	20th	1	122	0.64	0.63	0.21	0.84
		2	137	0.60	0.59	0.60	0.62
		3	134	0.69	0.64	0.53	0.68
	40th	1	122	0.78	0.48	0.53	0.63
		2	137	0.72	0.44	0.70	0.45
		3	134	0.60	0.60	0.60	0.60
NWF-CLS	20th	1	122	0.71	0.70	0.71	0.72
		2	137	0.77	0.69	0.77	0.71
		3	130	0.80	0.69	0.73	0.77
	40th	1	122	0.90	0.37	0.67	0.69
		2	137	0.86	0.44	0.82	0.52
		3	130	0.77	0.59	0.69	0.71
NWF-WRC	20th	1	122	0.75	0.60	0.64	0.67
		2	137	0.63	0.70	0.80	0.55
		3	130	0.73	0.68	0.77	0.61
	40th	1	122	0.95	0.40	0.83	0.56
		2	137	0.83	0.64	0.87	0.52
		3	130	0.77	0.55	0.75	0.61
WRF	20th	1	120	0.78	0.67	0.78	0.74
		2	137	0.83	0.73	0.83	0.73
		3	134	0.75	0.72	0.63	0.86
	40th	1	120	0.95	0.40	0.84	0.65
		2	137	0.82	0.65	0.80	0.68
		3	134	0.79	0.72	0.79	0.73

Table 2.3 Receiver Operating Curve Results for DIBELS 8 First Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
ORF	20th	1	113	0.77	0.72	0.73	0.77
		2	135	0.80	0.73	0.74	0.78
		3	132	0.75	0.81	0.66	0.83
	40th	1	113	0.94	0.41	0.80	0.73
		2	135	0.94	0.43	0.84	0.66
		3	132	0.78	0.74	0.69	0.77
ORF-ACC	20th	1	113	0.85	0.70	0.77	0.71
		2	135	0.71	0.77	0.71	0.78
		3	132	0.75	0.70	0.72	0.73
	40th	1	113	0.93	0.53	0.93	0.53
		2	135	0.93	0.40	0.93	0.42
		3	132	0.79	0.72	0.79	0.74

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year.

Table 2.4 Receiver Operating Curve Results for DIBELS 8 Second Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
NWF-CLS	20th	1	140	0.74	0.77	0.72	0.78
		2	193	0.80	0.80	0.75	0.85
		3	198	0.78	0.85	0.73	0.88
	40th	1	140	0.81	0.60	0.72	0.71
		2	193	0.80	0.76	0.80	0.74
		3	198	0.83	0.82	0.83	0.83
NWF-WRC	20th	1	140	0.80	0.73	0.70	0.72
		2	193	0.86	0.76	0.86	0.80
		3	198	0.86	0.80	0.80	0.82
	40th	1	140	0.85	0.50	0.85	0.54
		2	193	0.87	0.68	0.86	0.73
		3	198	0.85	0.78	0.85	0.82

Table 2.4 Receiver Operating Curve Results for DIBELS 8 Second Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
WRF	20th	1	140	0.83	0.81	0.78	0.85
		2	194	0.90	0.81	0.84	0.89
		3	198	0.90	0.83	0.82	0.92
	40th	1	140	0.82	0.68	0.78	0.79
		2	194	0.83	0.80	0.78	0.81
		3	198	0.86	0.81	0.83	0.85
ORF	20th	1	135	0.79	0.82	0.76	0.86
		2	193	0.86	0.82	0.86	0.83
		3	198	0.88	0.88	0.82	0.88
	40th	1	135	0.78	0.71	0.73	0.74
		2	193	0.83	0.79	0.83	0.75
		3	198	0.81	0.84	0.81	0.84
ORF-ACC	20th	1	135	0.74	0.86	0.74	0.86
		2	193	0.76	0.89	0.76	0.90
		3	198	0.86	0.71	0.55	0.94
	40th	1	135	0.75	0.71	0.71	0.79
		2	193	0.88	0.69	0.77	0.77
		3	NA	NA	NA	0.55	0.93
Maze	20th	1	189	0.94	0.76	0.72	0.84
		2	153	0.93	0.68	0.80	0.84
		3	190	0.94	0.66	0.81	0.83
	40th	1	189	0.97	0.41	0.82	0.82
		2	153	0.96	0.38	0.78	0.78
		3	190	0.94	0.36	0.81	0.85

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year. NA = Not applicable (i.e., no cut score for this period).

Table 2.5 Receiver Operating Curve Results for DIBELS 8 Third Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
NWF-CLS	20th	1	114	0.65	0.60	0.60	0.79
		2	172	0.68	0.69	0.59	0.74
		3	179	0.72	0.70	0.68	0.75
	40th	1	114	0.70	0.66	0.66	0.70
		2	172	0.79	0.74	0.70	0.79
		3	179	0.77	0.71	0.70	0.75
NWF-WRC	20th	1	114	0.65	0.62	0.65	0.64
		2	172	0.59	0.69	0.59	0.73
		3	179	0.68	0.70	0.64	0.72
	40th	1	114	0.79	0.57	0.74	0.61
		2	172	0.81	0.71	0.81	0.75
		3	179	0.74	0.74	0.74	0.74
WRF	20th	1	114	0.70	0.72	0.60	0.81
		2	172	0.73	0.73	0.68	0.79
		3	179	0.76	0.73	0.72	0.73
	40th	1	114	0.76	0.57	0.64	0.67
		2	172	0.79	0.68	0.72	0.77
		3	179	0.79	0.74	0.79	0.74
ORF	20th	1	114	0.70	0.70	0.65	0.73
		2	171	0.77	0.79	0.68	0.80
		3	179	0.72	0.84	0.72	0.84
	40th	1	114	0.77	0.51	0.68	0.64
		2	171	0.77	0.69	0.77	0.71
		3	179	0.79	0.76	0.79	0.74
ORF-ACC	20th	1	114	0.80	0.51	0.55	0.74
		2	171	0.86	0.48	0.36	0.92
		3	179	0.72	0.38	0.16	0.97
	40th	1	114	0.98	0.16	0.81	0.52
		2	171	NA	NA	0.47	0.86
		3	179	NA	NA	0.21	0.93

Table 2.5 Receiver Operating Curve Results for DIBELS 8 Third Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
Maze	20th	1	166	0.91	0.61	0.71	0.82
		2	168	0.95	0.61	0.76	0.78
		3	171	0.96	0.56	0.83	0.82
	40th	1	166	>.99	0.38	0.78	0.78
		2	168	>.99	0.28	0.75	0.79
		3	171	0.98	0.43	0.70	0.80

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year. NA = Not applicable (i.e., no cut score for this period).

Table 2.6 Receiver Operating Curve Results for DIBELS 8 Fourth Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
ORF	20th	1	134	0.70	0.81	0.70	0.83
		2	187	0.86	0.88	0.83	0.89
		3	189	0.75	0.84	0.72	0.85
	40th	1	134	0.66	0.76	0.66	0.78
		2	187	0.71	0.90	0.71	0.90
		3	189	0.67	0.80	0.71	0.80
ORF-ACC	20th	1	134	0.79	0.73	0.58	0.88
		2	187	0.89	0.37	0.31	0.99
		3	189	0.78	0.48	0.28	0.99
	40th	1	134	0.96	0.11	0.63	0.80
		2	187	NA	NA	0.22	0.98
		3	189	NA	NA	0.24	0.98
Maze	20th	1	182	0.97	0.72	0.78	0.86
		2	184	0.91	0.72	0.89	0.87
		3	181	0.94	0.67	0.94	0.76
	40th	1	182	0.99	0.43	0.71	0.80
		2	184	0.95	0.50	0.71	0.77
		3	181	0.90	0.49	0.76	0.72

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year.

Table 2.7 Receiver Operating Curve Results for DIBELS 8 Fifth Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
ORF	20th	1	96	0.71	0.76	0.71	0.79
		2	145	0.77	0.81	0.73	0.83
		3	149	0.70	0.83	0.78	0.78
	40th	1	96	0.78	0.65	0.78	0.67
		2	145	0.80	0.81	0.78	0.81
		3	149	0.79	0.78	0.83	0.72
ORF-ACC	20th	1	96	0.67	0.74	0.46	0.89
		2	145	>.99	0.40	0.31	0.98
		3	149	0.78	0.40	0.15	0.99
	40th	1	96	0.91	0.26	0.56	0.80
		2	149	NA	NA	0.30	0.96
		3	96	NA	NA	0.17	0.97
Maze	20th	1	142	0.84	0.50	0.76	0.74
		2	133	>.99	0.33	0.71	0.87
		3	109	0.93	0.42	0.64	0.89
	40th	1	142	0.92	0.29	0.75	0.66
		2	133	>.99	0.02	0.68	0.84
		3	109	>.99	0.22	0.64	0.88

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year. NA = Not applicable (i.e., no cut score for this period).

Table 2.8 Receiver Operating Curve Results for DIBELS 8 Sixth Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
ORF	20th	1	46	0.71	0.82	0.71	0.82
		2	153	0.79	0.84	0.79	0.87
		3	156	0.68	0.87	0.68	0.88
	40th	1	46	0.65	0.65	0.65	0.65
		2	153	0.70	0.86	0.75	0.79
		3	156	0.56	0.76	0.57	0.74

Table 2.8 Receiver Operating Curve Results for DIBELS 8 Sixth Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
ORF-ACC	20th	1	46	0.86	0.72	0.43	0.95
		2	153	>.99	0.41	0.26	0.98
		3	156	0.71	0.60	0.08	0.98
	40th	1	46	0.95	0.23	0.40	0.85
		2	153	NA	NA	0.35	0.92
		3	156	NA	NA	0.14	0.97
Maze	20th	1	101	>.99	0.62	0.86	0.87
		2	103	>.99	0.72	0.78	0.95
		3	104	0.89	0.67	0.78	0.88
	40th	1	101	>.99	0.33	0.72	0.83
		2	103	0.96	0.59	0.59	0.93
		3	104	0.96	0.60	0.78	0.86

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year. NA = Not applicable (i.e., no cut score for this period).

Table 2.9 Receiver Operating Curve Results for DIBELS 8 Seventh Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
ORF	20th	1	37	0.91	0.96	0.91	0.96
		2	155	0.79	0.80	0.79	0.81
		3	155	0.74	0.88	0.79	0.80
	40th	1	37	0.79	0.78	0.79	0.78
		2	155	0.84	0.72	0.87	0.69
		3	155	0.74	0.79	0.80	0.76
ORF-ACC	20th	1	37	>.99	0.35	0.18	>.99
		2	155	0.93	0.49	0.07	0.97
		3	155	0.76	0.59	0.07	0.99
	40th	1	37	NA	NA	0.32	>.99
		2	155	NA	NA	0.39	0.93
		3	155	NA	NA	0.20	0.98

Table 2.9 Receiver Operating Curve Results for DIBELS 8 Seventh Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
Maze	20th	1	93	>.99	0.81	0.70	0.90
		2	93	>.99	0.74	0.92	0.89
		3	95	0.92	0.77	0.85	0.85
	40th	1	93	>.99	0.50	0.76	0.86
		2	93	>.99	0.63	0.83	0.80
		3	95	>.99	0.35	0.88	0.86

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year. NA = Not applicable (i.e., no cut score for this period).

Table 2.10 Receiver Operating Curve Results for DIBELS 8 Eighth Grade Subtests Predicting End-of-year Iowa Total Reading Scores

Measure	Criterion	Period	N	Previous		Updated	
				Sensitivity	Specificity	Sensitivity	Specificity
ORF	20th	1	45	0.81	0.79	0.71	0.83
		2	105	0.86	0.74	0.84	0.77
		3	112	0.75	0.63	0.76	0.67
	40th	1	45	0.78	0.77	0.78	0.77
		2	105	0.78	0.63	0.84	0.55
		3	112	0.75	0.55	0.75	0.55
ORF–ACC	20th	1	45	0.91	0.08	0.24	>.99
		2	105	0.96	0.10	0.18	0.98
		3	112	0.96	0.13	0.13	0.97
	40th	1	45	NA	NA	0.44	0.92
		2	105	NA	NA	0.34	0.95
		3	112	NA	NA	0.32	0.93
Maze	20th	1	46	0.91	0.63	0.73	0.88
		2	50	0.92	0.65	0.63	0.81
		3	52	0.83	0.64	0.71	0.82
	40th	1	46	0.91	0.23	0.82	0.85
		2	50	0.97	0.14	0.75	0.61
		3	52	0.97	0.27	0.70	0.73

Note. Criteria were percentile ranks. 1 = Beginning of year. 2 = Middle of year. 3 = End of year. NA = Not applicable (i.e., no cut score for this period).

References

References

Betts, E.A. (1946). Foundations of reading instruction, with emphasis on differentiated guidance. New York: American Book

Halladay, J. L. (2012). Revisiting key assumptions of the reading level framework. The Reading Teacher, 66(1), 53-62.

Treptow, M. A., Burns, M. K., & McComas, J. J. (2007). Reading at the frustration, instructional, and independent levels: The effects on students' reading comprehension and time on task. School Psychology Review, 36(1), 159-166.