

Dynamic Indicators of Basic Early Literacy Skills

8th Edition

Administration and Scoring Guide 2026 Edition

University of Oregon (2026). 8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS®): Administration and Scoring Guide, 2026 Edition. Eugene, OR: University of Oregon. Available: https://dibels.uoregon.edu

Contributing Authors

Gina Biancarosa, EdD
Associate Professor and Ann Swindells Chair in Education
Department of Special Education and Clinical Sciences,
College of Education, University of Oregon

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

Sunhi Park
Graduate Research Assistant Fellow
Center on Teaching & Learning
College of Education, University of Oregon

Janet Otterstedt, MS
Research Assistant and Project Coordinator
Center on Teaching & Learning
College of Education, University of Oregon

Contributing Editors

Maureen Warman, MS
Senior Research Assistant I and DIBELS Data System Manager
Center on Teaching & Learning
College of Education, University of Oregon

David Larsen, MA
Senior Research Assistant II and DIBELS Data System Customer Support Manager
Center on Teaching & Learning
College of Education, University of Oregon

Anna D. Ingram
Administrative Program Assistant
Center on Teaching & Learning
College of Education, University of Oregon

University of Oregon would also like to acknowledge and thank our partner Amplify Education for their work on copy editing and formatting.

Table of Contents

Chapter 1: Introduction to DIBELS® 8th Edition 6

DIBELS History 6

Dimensions of Reading Assessed by DIBELS 8 7

Description of DIBELS 8 8

Appropriate Uses of DIBELS 8 11

Rationale for and Innovations in DIBELS 8 12

Development of DIBELS 8 21

Summary 35

Chapter 2: Administration Instructions and Scoring Procedures 36

DIBELS 8th Edition General Guidelines 36

Important Considerations for Inclusive and Accurate Scoring 43

Letter Naming Fluency (LNF) 49

Phonemic Segmentation Fluency (PSF) 53

Nonsense Word Fluency (NWF) 61

Word Reading Fluency (WRF) 70

Oral Reading Fluency (ORF) 73

Maze 78

Chapter 3: Interpreting DIBELS 8th Edition Scores 86

DIBELS 8 Composite Scores 86

DIBELS 8 Risk Classification 87

DIBELS 8 Percentile Ranks 88

DIBELS 8 Raw scores 88

DIBELS 8 Zones of Growth® 89

DIBELS 8th Edition and Dyslexia Screening 89

Cautions in Interpreting DIBELS 8 Scores 90

Chapter 4: Progress Monitoring with DIBELS 8th Edition 92

Choosing a Subtest for Progress Monitoring 92

Frequency of Progress Monitoring 93

Determining Response to Intervention 94

Off-grade Progress Monitoring 95

References 97

Appendix A: ORF Passage Statistics 107

Appendix B: Maze Passage Statistics 115

Appendix C: DIBELS 8th Edition Pronunciation Guide 122

Appendix D: Administration and Scoring Fidelity Checklists 123
Letter Naming Fluency Fidelity Checklist 123
Nonsense Word Fluency Fidelity Checklist 125
Word Reading Fluency Fidelity Checklist 126
Oral Reading Fluency Fidelity Checklist 127
Maze Fidelity Checklist 128
Appendix E: DIBELS 8th Edition Benchmark Goals 130
Appendix F: Composite Score Derivation and Calculation Guide 136
Derivation of the DIBELS 8 Composite Score 136
Calculating the DIBELS 8 Composite Score 138
Example with Full Data 143
Example with Discontinue Rule Implemented in Kindergarten 144
Example with Gating Rule Implemented in Second Grade 145
DIBELS 8th Edition Composite Score Calculation Worksheet Kindergarten 146
DIBELS 8th Edition Composite Score Calculation Worksheet First Grade 147
DIBELS 8th Edition Composite Score Calculation Worksheet Second Grade 148
DIBELS 8th Edition Composite Score Calculation Worksheet Third Grade 149
DIBELS 8th Edition Composite Score Calculation Worksheet Fourth Grade 150
DIBELS 8th Edition Composite Score Calculation Worksheet Fifth Grade 151
DIBELS 8th Edition Composite Score Calculation Worksheet Sixth Grade 152
DIBELS 8th Edition Composite Score Calculation Worksheet Seventh Grade 153
DIBELS 8th Edition Composite Score Calculation Worksheet Eighth Grade 154
Index 155

Chapter 1: Introduction to DIBELS® 8th Edition

This manual is a compendium of information regarding DIBELS 8th Edition. It details the nature and purpose of DIBELS 8, how DIBELS 8 differs from previous editions of DIBELS, how to administer and score DIBELS 8 benchmark subtests three times per year (Beginning, Middle, and End), and when and how to use benchmark, and progress monitoring data to inform instructional decision-making. It also provides appendices for passage statistics, fidelity of implementation checklists, and composite scores. This manual begins with an overview of the history of the DIBELS assessment system.

DIBELS History

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) consists of a set of measures for assessing reading skills. DIBELS began as a series of short tests that assessed early childhood literacy in kindergarten and first grade (Meyer, 2000). Over the years, DIBELS has gone through several editions, expanding the range of skills assessed and grades in which it can be used. DIBELS is now in its 8th Edition, which offers reading measures for Grades K-8.

DIBELS began as Dynamic Indicators of Basic Skills (DIBS; Shinn, 1989, 1998). Inspired by Deno's (1986) definition of curriculum-based measurement (CBM), DIBS, and DIBELS after it, was an attempt to ground classroom assessment practices and decision making in measurement science. With the support of a federal grant, the first DIBELS measures intended for use in the elementary grades (i.e., kindergarten and first grade) were developed as part of Dr. Ruth Kaminski's doctoral thesis in 1992 at the University of Oregon, where Dr. Roland Good served as her advisor. The measures were Letter Naming Fluency, Picture Naming Fluency, and Phonemic Segmentation Fluency. In the years since, the evolution of DIBELS measures and their interpretation has involved a number of University of Oregon faculty in addition to Dr. Good and Dr. Kaminski, including, but not limited to, Dr. Edward Kame'enui, Dr. Mark Shinn, and Dr. Deborah Simmons. In addition, numerous University of Oregon graduate students have contributed to the rich history of DIBELS research and development, including Dr. Sylvia Barnes Smith, Dr. Rebecca Briggs, Dr. Kelli Cummings, Dr. Deborah Laimon, and Dr. Kelly Powell-Smith, among others.

Updated editions of DIBELS have been released every several years beginning in 1996. Before

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition, the last update (DIBELS Next) was in 2010 and before that in 2002 (DIBELS 6th Edition). Over the years, subtests have come (e.g., Nonsense Word Reading Fluency, Oral Reading Fluency) and gone (e.g., Picture Naming Fluency, Initial Sound Fluency). DIBELS 8th Edition continues the legacy of development and research that has been ongoing at the University of Oregon since the late 1980s. It introduces several changes, including new features such as measures spanning kindergarten through eighth grade, a new DIBELS measure (Word Reading Fluency), and modern measurement approaches to scoring, as well as the retirement of two existing measures (First Sound Fluency and Retell Fluency).

Dimensions of Reading Assessed by DIBELS 8

DIBELS 8th Edition offers six subtests designed to assess component skills involved in reading: Letter Naming Fluency (LNF), Phonemic Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), Word Reading Fluency (WRF), Oral Reading Fluency (ORF), and Maze. These subtests are aligned to four of the five "Big Ideas" in reading identified by the National Reading Panel (National Institute of Child Health and Human Development, 2000), including phonological awareness (which encompasses phonemic awareness), phonics (or the alphabetic principle), fluency, and comprehension (Riedel, 2007; see Table 1.1). In many ways the DIBELS subtests represent not only the constructs in the National Reading Panel Report (NICHD, 2000), but also a developmental continuum. As a result, the subtests included change across grades in a manner that parallels student development and instructional foci (Adams, 1990; Chall, 1996; Ehri, 2005; Paris & Hamilton, 2009).

Table 1.1 The Big Ideas in Reading and DIBELS 8 Subtests

Big Idea	LNF	PSF	NWF	WRF	ORF	Maze
Phonemic awareness		X				
Phonics			X	X	X	
Accuracy and fluency with text				X	X	X
Vocabulary						
Comprehension					X	X

©2026 University of Oregon. All rights reserved.

Some DIBELS 8 subtests are also aligned to subskills of reading that are associated with risk for dyslexia and other word reading disabilities. The International Dyslexia Association (IDA) recommends universal screening of students in kindergarten through second grade (IDA, 2020). Consistent with IDA recommendations, DIBELS 8 offers LNF, PSF, and NWF subtests as dyslexia screening measures of rapid naming (or processing speed), phonemic awareness, and letter-sound correspondence for use in kindergarten and first grade. Also consistent with IDA recommendations, DIBELS 8 offers real and nonsense word measures (NWF, WRF, and ORF) as dyslexia screening measures.

Description of DIBELS 8

DIBELS 8th Edition takes a curriculum-based measurement (CBM) approach to assessing reading. It is intended for assessing reading skills from the beginning of kindergarten through the end of eighth grade. DIBELS 8 subtests are designed as brief, easily administered measures of reading. Five of the subtests (LNF, PSF, NWF, WRF, and ORF) are 60-second measures designed to be administered individually in a quiet setting. The sixth subtest, Maze, is a 3-minute measure designed to be administered in group settings. Because DIBELS subtests are timed measures, efficiency in reading skills is considered as well as accuracy. The subtests offered in specific grades are aligned to curriculum and instruction typical for each grade, as well as to recommendations made by the IDA (see Figure 1.1).

Kindergarten	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
LNF	LNF							
PSF	PSF							
NWF	NWF	NWF	NWF					
WRF	WRF	WRF	WRF					
	ORF	ORF	ORF	ORF	ORF	ORF	ORF	ORF
		MAZE	MAZE	MAZE	MAZE	MAZE	MAZE	MAZE
Kindergarten	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8

Figure 1.1. DIBELS 8th Edition Timeline of Subtest Availability by Grade

©2026 University of Oregon. All rights reserved.

To maintain efficiency of benchmark assessment procedures, we have instituted new discontinuation rules to save time and avoid student frustration during benchmark assessment. As a result, total administration time varies by grade and by student skill (see Table 1.2).

Table 1.2 Administration Time in Minutes by Grade and Administration Type

Administration Type	K	1	2-3	4-8
Individual	4-6	5-7	4	2
Group	NA	NA	5	5

Note. Ranges are provided in grades where rules exist for discontinuing a benchmark assessment. Only Maze is administered in a group setting. NA = not applicable.

Letter Naming Fluency (LNF). LNF is a standardized, individually-administered test that provides a measure of risk for reading achievement. LNF is based on research by Marston and Magnusson (1988) and is administered to students in the beginning of kindergarten through the end of first grade.

For LNF, students are presented with a page of 100 uppercase and lowercase letters arranged in a random order and are asked to name as many letters as they can. Students are given one minute to provide letter names. If a student does not know a letter name, the examiner provides the letter name and marks the letter name incorrect. The LNF measure has three benchmark forms for each grade in which it is available. As in previous editions, alternate progress-monitoring forms are not provided for LNF because it serves solely as a risk indicator.

Phonemic Segmentation Fluency (PSF). PSF is a standardized, individually-administered measure of phonological awareness. PSF is a good predictor of reading achievement and is administered to students in the beginning of kindergarten through the end of first grade.

PSF assesses students' ability to fluently segment two- to six-phoneme words into their individual phonemes. In PSF, the examiner orally presents a series of words and asks a student to verbally produce the individual phonemes for each word. For example, if the examiner said "sat," and the student said "/s/ /a/ /t/", the student would receive three points for the word. After each response, the examiner presents the next word. Students are given one minute to segment the words into phonemes. The PSF measure has three benchmark forms and 20 alternate progress-monitoring

©2026 University of Oregon. All rights reserved.

forms for each grade in which it is available.

Nonsense Word Fluency (NWF). NWF is a standardized, individually-administered measure of the alphabetic principle. NWF is seen as a "pure" measure of the alphabetic principle, because vocabulary and sight word knowledge cannot play a role in recognizing nonsense words. NWF is administered to students in the beginning of kindergarten through the end of third grade.

NWF assesses students' ability to decode words based on the alphabetic principle. For NWF, students are presented with an 8.5-inch x 11-inch sheet of paper with nonsense words (e.g., sig, ral) and asked to verbally produce (a) the whole nonsense word or (b) individual letter sounds. For example, if the stimulus word is "hap", a student could say the nonsense word as a whole or "/h/ /a/ /p/" to receive three letter sounds correct. On DIBELS 6th Edition, if the nonsense word was read as a whole (either initially or after sounding out), the student received credit for one whole word read correctly. On DIBELS Next, the student only received credit for reading the nonsense word correctly if it was read as a whole in the initial attempt. DIBELS 8th Edition reverts to the DIBELS 6th Edition practice because it more accurately captures students' knowledge of sound-spelling patterns and the ability to blend sounds into words, which is the primary intent of NWF. Students are given one minute to read or sound out as many nonsense words as they can. The NWF measure has three benchmark forms and 20 alternate progress-monitoring forms for each grade in which it is available.

Word Reading Fluency (WRF). WRF is standardized, individually-administered measure of accuracy and fluency with lists of words. WRF is administered to students in the beginning of kindergarten through the end of third grade.

The new WRF subtest involves reading real words out of context. Inspired by other CBMs that incorporate WRF, most especially easyCBM (Alonzo & Tindal, 2007), it is a standardized, individually-administered measure of accuracy and fluency in reading "sight" words. Sight words include words with irregular pronunciations (non-decodable words like "the" and "was" and "of") as well as common words with regular pronunciations (decodable words like "in" and "we" and "no"). WRF is administered to students from the beginning of kindergarten through the end of third grade.

In WRF, students are presented with an 8.5-inch x 11-inch sheet of paper with real words and asked to verbally produce the whole word. Students must blend words to receive credit. In contrast to

©2026 University of Oregon. All rights reserved.

NWF, no credit is given for individual letter sounds. Students are given one minute to read as many words as they can, and the final score is the number of words read correctly within one minute. The WRF measure has three benchmark forms and 20 alternate progress-monitoring forms for each grade in which it is available.

Oral Reading Fluency (ORF). ORF is a standardized, individually-administered measure of accuracy and fluency with connected text. ORF is administered to students in the beginning of first grade through the end of eighth grade.

ORF assesses a student's ability to read words in connected text. In ORF, the examiner presents the student with a passage and asks the student to read the passage aloud for one minute. Words omitted, substituted, and hesitations of more than three seconds are scored as errors. Words self-corrected within three seconds are scored as accurate. The final score is the number of words read correctly (and self-corrected) within one minute. The ORF measure has three benchmark forms and 20 alternate progress-monitoring forms for each grade in which it is available.

Maze. Maze is a standardized, group-administered measure of reading comprehension. Maze is administered to students in the beginning of second grade through the end of eighth grade. In Maze, the examiner presents students with a passage that has every seventh word removed and replaced with three options. In third through eighth grade, the first and last sentence are left intact, and in second grade, the first two sentences and last sentence are left intact. The final score is one-half the number of overt errors subtracted from the number of maze words selected correctly within three minutes. Skipped items are treated as errors, but items not reached are not counted as errors. The Maze measure has three benchmark forms and several progress monitoring forms available for second through eighth grade. We offer fewer Maze progress monitoring forms than for other subtests because we do not recommend more than monthly progress monitoring for Maze.

Appropriate Uses of DIBELS 8

DIBELS 8 measures are designed to be used from the beginning of kindergarten through the end of eighth grade. Although DIBELS 8 can be used for off-grade assessment, it has not been validated for this use.

©2026 University of Oregon. All rights reserved.

DIBELS 8 subtests were developed and researched as indicators of risk and progress in overall reading, as well as risk for dyslexia and other reading difficulties. DIBELS 8 has three principal uses: to identify students who may be at risk of reading difficulties by screening up to three times per year, to document students' progress of reading skills as a consequence of special intervention programs through progress monitoring, and to provide minimum levels of performance for all students to reach to be considered on track for becoming a reader through benchmark goals and timelines. DIBELS 8 benchmark forms were validated as screening measures administered at the beginning, middle, and end of a school year. Additional forms have been validated for use in progress monitoring and are provided for most measures.

DIBELS was also designed as a means to evaluate the effectiveness of intervention for those students receiving support, in order to inform changes in intervention strategy as necessary to improve student learning and growth. Similarly, DIBELS was designed for use in research on reading development, especially the development of readers at risk.

DIBELS 8 can be used to make judgments about the instructional needs and responsiveness of individual students and regarding the efficacy of curriculum and instructional practices. It has not been designed to render judgments regarding teacher effectiveness or school progress. See chapters 3 and 4 for guidance on how to interpret DIBELS 8th Edition scores for screening and progress monitoring.

Rationale for and Innovations in DIBELS 8

DIBELS 8 was developed consistent with best practices in educational measurement (AERA, APA, & NCME, 2014). The most recent standards for educational testing suggest that test creators must renorm tests "with sufficient frequency to permit continued accurate and appropriate score interpretations" (AERA et al., 2014, p. 104). This standard is interpreted as meaning that not only should norms be updated regularly (as DIBELS Next norms have been), but related validity evidence must also be updated, especially when used for critical instructional decision-making. Given that validity data on the last edition of DIBELS was almost a decade old and educational practices have shifted during that decade (for example, the introduction of Common Core Standards and a new generation of state tests), a new edition of DIBELS was deemed necessary. In addition, a new edition provided an opportunity to improve DIBELS in several ways. These innovations are summarized in this section.

©2026 University of Oregon. All rights reserved.

Expanded grade levels. DIBELS 8 can be used in kindergarten through eighth grade. The expansion through eighth grade means that DIBELS can now be used in schools with a wider range of grade configurations: K-3, K-5, K-8, 5-8, 6-8, etc.

Consistent subtests within grade. DIBELS 8 subtests used at any point during a given grade are available for all benchmark periods in that grade (see Figure 1.1 on page 8). This availability supports users who may want or be required to have consistent data across all three benchmark assessment periods in the school year.

Discontinue and gating benchmark rules. To maintain efficiency of benchmark assessment procedures, we have instituted new discontinuation rules in kindergarten and first grade (see Chapter 2 on page 36 for more details). These rules are intended to save time and avoid student frustration during benchmark assessment. For example, in the beginning of the school year, if a kindergarten student cannot segment any phonemes on PSF, or if a first grade student cannot read any words on WRF, the administrator does not need to administer the remaining subtests (NWF and WRF in kindergarten and ORF in first grade). In this way, administration remains efficient, while still yielding information on more able readers.

These rules are intended to save time and to spare the student unnecessary frustration. The discontinue benchmark rules were derived from a national field trial that indicated students who scored 0 for specific assessments at specific benchmark periods were extremely unlikely to get any items correct on the remaining subtests. The gating benchmark rules were derived from the same study, which demonstrated that students scoring well above benchmark for specific assessments in specific grades were extremely unlikely to demonstrate any risk on the remaining assessments. Nonetheless, examiners have the option of administering the remaining subtests based on professional judgment. See Chapter 2 on page 36 for more details on when to discontinue and gate students and what values to use when computing composites scores for discontinued and gated students.

Font type and size. The font chosen for DIBELS 8th Edition was informed by research on the effect of fonts for children with and without word reading disabilities like dyslexia. Although a great deal of research has explored the effects of different fonts, including “dyslexia friendly” fonts like Dyslexie and Open Dyslexic, very few of these studies used rigorous scientific methods. The few peer-reviewed

©2026 University of Oregon. All rights reserved.

studies that have employed randomized trials have yielded equivocal results. Dyslexia-friendly fonts have no discernible effects on readers with and without dyslexia and other word reading disabilities. For example, children with and without dyslexia showed no significant differences in reading speed or reading accuracy when the Dyslexie, Times New Roman, and Arial fonts were compared, especially when spacing of letters was controlled across fonts (Duranovic, Senka, & Babic-Gavric, 2018; Marinus et al., 2016; Wery & Diliberto, 2017). More promising is research that shows that the spacing of letters, which co-varies with font-size, does affect reading speed and comprehension for all readers.

As a result, we paid a great deal of attention to font sizes in the development of DIBELS 8th Edition (see next section on page 21), but the ultimate choice of font was guided by the distinguishability of letters. Of paramount concern was that the capital i (I) be easily distinguishable from a lowercase L (l). To accomplish this aim, it was necessary to use a font with serifs, which are the slight projections on letters in some fonts. For example, in Arial font, which does not have serifs, the uppercase i and lowercase L are nearly indistinguishable: I, l.

However, no serif font represents the letters a, g, j, and q in the forms they are more commonly taught in the primary grades: *a*, ɑ, g, j, and *q*. In fact, any font that represents one of these letters as they are typically taught represents other letters in a less typical form. Given that no font fulfilled all of these practical considerations, we opted for the Rockwell font, which is similar to the more familiar Times New Roman but has slightly thicker serifs and a more typical form of lower-case G (see Table 1.3 page 14). DIBELS PDF forms are locked so font types cannot be changed. Changing the font type invalidates the measures. However, the font size can be changed as an accommodation for students with low vision.

Table 1.3 DIBELS 8th Edition Fonts

Times New Roman	Rockwell
i, I, l, L	i, I, l, L
a, g, j, q	a, g, j, q

Rockwell is used for all subtests except for Oral Reading Fluency (ORF) and Maze, where Times New Roman is used instead. Times New Roman is used when subtest probes involve reading

©2026 University of Oregon. All rights reserved.

in context because research has shown that young readers, including those with dyslexia and other word reading difficulties, prefer familiar fonts (Kuster, van Weerdenburg, Gompel, & Bosman, 2018; Wery & Diliberto, 2017). This same research indicated that font and preference did not affect overall performance for any group of readers. As a result, we used the more familiar Times New Roman to minimize any potential interference in meaning-making that a less familiar font might cause.

Font sizes for DIBELS 8th Edition were informed by research on the effects of font sizes for children with and without word reading disabilities like dyslexia. For all readers, larger font sizes promote faster reading speeds up to a "critical" font size when increases in font size no longer result in faster reading (O'Brien, Mansfield, & Legge, 2005). This critical font decreases in size with grade level for all readers, suggesting that font sizes can be safely decreased each year for all readers. Although readers with dyslexia benefit from larger font sizes than readers without dyslexia, their ability to read smaller font each year efficiently progresses in a similar fashion. Additional research demonstrates that smaller font sizes and longer line lengths can also interfere with primary grade readers' ability to comprehend text (Katzir, Hershko, & Halamish, 2013). However, for intermediate grade readers larger font sizes interfered with comprehension, while line lengths had no effect. Spacing between lines had no effect for either group of readers. In general, research indicates that larger print results in younger readers reading faster and comprehending better, but that there are diminishing and even no benefits the older a reader is (Hughes & Wilkins, 2000; Katzir et al., 2013; O'Brien et al., 2005; Wilkins, Cleave, Grayson, & Wilson, 2009). Thus, DIBELS 8th Edition font sizes start at 24pt in Kindergarten and get slowly, but progressively smaller until fifth grade (see Table 1.4 on page 16). ORF font sizes are slightly smaller than font sizes for the other DIBELS 8th Edition subtests to keep passages from taking up more than the front and back of a single page without resorting to overly narrow margins.

©2026 University of Oregon. All rights reserved.

Table 1.4 DIBELS 8th Edition Font Sizes

Subtest	K	1	2	3	4	5+
LNF	24pt	24pt	NA	NA	NA	NA
NWF	24pt	22pt	20pt	18pt	NA	NA
WRF	24pt	22pt	20pt	18pt	NA	NA
ORF	NA	20pt	18pt	16pt	14pt	13pt
Maze	NA	NA	18pt	16pt	14pt	13pt

Note. NA = not applicable.

Letter Naming Fluency improvements. For DIBELS 8th Edition, LNF now accounts for how frequently letters appear in both uppercase and lowercase forms. To better control differences in difficulty between forms, consistent rules are used in both kindergarten and first grade regarding when less frequent letters can appear on the forms. Each form in both grades begins with a sampling of the 20 most frequently seen letters (Jones & Mewhort, 2004), thereby preventing students from getting frustrated by forms that begin with rarer letters, such as X or q. The kindergarten version of LNF also only assesses the 40 most commonly seen uppercase and lowercase letters, while the first grade version assesses 49 uppercase and lowercase letters.

LNF *excludes* three letters on all forms: uppercase and lowercase W and lowercase L. Although these are obviously important letters for students to know, they introduce real problems in a fluency assessment. W is the only letter with a multi-syllabic name: three syllables to be exact. As a result, any time W appears, it takes three times as long to name as other letters, which negatively affects a student's LNF score. The lowercase L (l) was eliminated because it is easily confused with both the uppercase I and the number 1. Not only does this visual similarity pose problems for students, but it has also historically created

©2026 University of Oregon. All rights reserved.

scoring problems for the adult administering the assessment. By avoiding these letters, each included item (or letter) is equally challenging, other than in terms of its frequency in printed language.

Phonemic Segmentation Fluency improvements. In DIBELS 8th Edition, PSF accounts for both word frequency and the number of phonemes in a word. All forms draw only from the 2,500 most frequent words in English (Balota et al., 2007) to minimize vocabulary familiarity from interfering with student performance. In addition, to better control differences in difficulty between forms, consistent rules are used in both grades regarding where less frequent words can appear on the forms. Moreover, spelling patterns are ordered in terms of the number of phonemes, proceeding from two phoneme words to words with progressively more phonemes.

In kindergarten, the first 20% of items have two phonemes, while the remaining 80% have three phonemes. In this way, PSF now reduces the distinct floor effects (i.e., many students scoring zero) in kindergarten that have plagued previous versions and, thus, eliminates the need for a separate measure of initial sound fluency. In first grade, the progression in difficulty is a bit more rapid, with the first 13% of items having two phonemes and then increasing in phonemes with additional increases after every eight items.

Nonsense Word Fluency improvements. In DIBELS 8th Edition, NWF now accounts for the frequency of spelling patterns (Jones & Mewhort, 2004; Norvig, 2012). As a result, all forms utilize only phonetically regular letter combinations that actually appear in English. Thus, students will no longer be asked to decode nonsense words like “fev” or “kaj”, and nonsense words like “kex” will appear less often than ones like “lat”.

DIBELS 8th Edition also expands the spelling patterns assessed beyond simply consonant-vowel-consonant (CVC) after kindergarten. While kindergarten forms are limited to CVC patterns, the first-grade forms also include vowel-consonant (VC) spelling patterns. In addition, the latter half of first grade forms include additional spelling patterns typically taught in first grade, thus increasing the instructional relevance of this DIBELS subtest. DIBELS 8th Edition also now offers NWF in second and third grade by including more complex phonics patterns in these grades. As a result, DIBELS NWF forms provide instructionally relevant information even for students who are at minimal risk in kindergarten through third grade. New spelling patterns included in first through third grade appear in

©2026 University of Oregon. All rights reserved.

Table 1.5 below.

Table 1.5 Examples of First through Third Grade NWF Spelling Patterns

Pattern	Grade introduced	Example non-word
CVCe	1	bace
CVr(C)	1	zart
CVCC	1	melb
CCVC	1	scap
CCVCC	1	brold
(C)CVVC(C)	2	geap
CVCCy	2	foddy
(C)V|CVC(C)	3	cotalm
(C)VC|CVC(C)	3	fudlerk

An additional improvement to NWF is that we have reverted to scoring words recoded correctly (WRC; DIBELS 6th Edition practice) rather than whole words read (WWR; DIBELS Next practice). Whereas with WWR students only received credit if they correctly read a nonsense word at first sight (i.e., without sounding out), with WRC they also receive credit if they blend a nonsense word after sounding out the component sounds. Because both methods of scoring predict student risk, in DIBELS 8th Edition, students receive credit for blending nonsense words whether they sound them out first or not. In addition, with WWR the information about students' ability to blend words was lost if students first verbalized the sounds prior to blending them into words. In this case students would receive no credit for whole words read even though they blended sounds into words. Since the main purpose of NWF is

©2026 University of Oregon. All rights reserved.

to assess readers' understanding of the alphabetic principle and sound-symbol correspondence, WRC was deemed the more appropriate scoring method.

Word Reading Fluency innovative features. WRF targets real words based on age of acquisition in students' vocabulary (Brysbaert & Biemiller, 2017) and their frequency in written text (Balota et al., 2007). WRF assesses only words that are typically acquired orally in or before a given grade. This reduces the likelihood that students will encounter words on the assessment that they have never heard before and are not yet expected to know.

In addition, each form starts with a sample of the most frequent words seen in text and then moves on to less frequent words in the latter half of the form. In this way, WRF yields instructionally relevant information both for students at risk and students at minimal risk.

Finally, DIBELS WRF accounts for word complexity, as measured by the number of syllables in a word. All forms include one-syllable words. Grades 1-3 include two-syllable words, and Grades 2-3 include two-syllable and three-syllable words. In Grade 3, we also included words with more than three syllables, but again only those that are typically acquired by Grade 3 and are frequently seen in print.

These features ensure the instructional relevance of DIBELS WRF results for all students. Importantly, our research, as well as that of others (Clemens, Shapiro, & Thoemmes, 2011; Fuchs, Fuchs, & Compton, 2004; Smith, Cummings, Nese, Alonzo, Fien, & Baker, 2014), has shown that the inclusion of WRF helps to identify students at risk who might otherwise be missed by other DIBELS subtests.

Oral Reading Fluency improvements. DIBELS 8th Edition marks the first time that DIBELS ORF requires the administration of only one passage per benchmark period. Research has shown that administering more than one passage does little to improve the reliability and validity of ORF, meaning that the minimal benefits of administering three passages just does not warrant the additional administration time (Baker et al., 2015; Petscher & Kim, 2011). Rather, a single passage works just as well, and reduces the testing burden for both students and assessors.

An additional unique and exciting feature of DIBELS 8th Edition ORF passages is that they were written by experienced and aspiring children's authors, most of whom have previous experience writing for students and have previously published short stories. The authors have diverse backgrounds, come

©2026 University of Oregon. All rights reserved.

from across the US, and have experience writing in a range of genres. As a result, ORF passages are not only more engaging for both students and assessors, but also read as more authentic and appropriate for the grades in which they appear.

Maze improvements. Maze has now been informed by research that shows consistently that maze measures tend to assess low-level comprehension (e.g., January & Ardoin, 2012; Shanahan, Kamil, & Tobin, 1982). To make DIBELS maze measures more informative, we undertook several innovations. First, as with ORF, maze passages are written by experienced and aspiring authors. Second, more work has gone into the selection of distractors, and this work is described in the section on development on page 21. Third, the formatting of Maze was revised to make reading the passages easier on the eye, reflecting research that suggests that overly long lines can cause disfluency and interfere with reading comprehension for young readers (e.g., Dyson & Haselgrove, 2001; Katzir et al., 2013). Finally, maze measures are available in second through eighth grade instead of only third through sixth.

Retirement of subtests. Both First Sound Fluency (FSF) and Retell Fluency (RTF) have been removed from DIBELS 8 as subtests for several reasons. First and most critically, both subtests add time to the administration of DIBELS without adding much useful information for screening or instructional planning. Thus, the information yielded through these measures relative to the time spent administering them was not deemed as worthwhile as it was for the other DIBELS subtests. Additional factors that played into the decision to drop FSF were its redundancy and constrained nature. Given our modifications to PSF, FSF was deemed more redundant with PSF than it had been in the past. In addition, First Sound Fluency taps a very constrained aspect of phonemic awareness: the detection of initial phonemes. Given that this phonemic awareness skill is mastered quite quickly, especially in the presence of instructional intervention, FSF was not a good candidate for administering in all three benchmark periods in kindergarten. In other words, its best use was incompatible with the new design specifications of DIBELS 8, which requires the same subtests to be available throughout a grade.

Additional factors that played into the decision to drop RTF were questions regarding its validity and the new availability of Maze in lower grades. DIBELS users have often questioned whether a words-per-minute rate for retelling captures comprehension adequately. When measured as a rate, factors unrelated to comprehension can radically affect scores. For example, students with

©2026 University of Oregon. All rights reserved.

speech impediments like stutters will produce fewer words in their retell regardless of their level of comprehension. Similarly, students learning English who have more limited expressive vocabulary than receptive vocabulary also have a tendency to score lower than their English-only speaking counterparts regardless of their level of comprehension. In other words, RTF was as much a measure of expressive language fluency as it was of reading comprehension. The influence of expressive language fluency is undesirable in a measure of reading comprehension.

Furthermore, recent research has demonstrated that retell in the absence of word reading fluency is almost useless. Word reading accounts for almost all the variability in first grade reading comprehension measures (e.g., Lonigan & Burgess, 2017; Lonigan, Burgess, & Schatschneider, 2018). In fact, word reading sets a hard limit on whether a student can read enough text to build a mental representation of what is read. For students who score below the risk cut-score in first grade, which is four or fewer words at the beginning of the year and 27 or fewer words at the end of the year, very little of substance has been read. As a result, only the students with the most advanced reading skills will be able to give a retelling that yields any useful information. However, this picture begins to change quickly past first grade. Consequently, Maze, which has superior predictive powers to retell fluency, is now available from Grade 2 onward.

Development of DIBELS 8

DIBELS 8 was developed consistent with the most recent standards in educational measurement (AERA et al., 2014). These standards provide criteria for test development that promote the validity of interpretations of test scores. The development process outlined by the standards includes detailing the intended uses of a test, specifying content and format requirements, and using specifications to create item pools, inform item selection, and guide assignment of items to forms. These standards also recommend an iterative approach to development decisions and evidence gathering.

Consistent with these standards, the intended uses of DIBELS 8 were defined (as noted in the section of this manual on Appropriate Uses on page 11). Prior to determining specifications, DIBELS researchers performed a comprehensive literature review of critiques and limitations of DIBELS and other reading CBMs. Researchers also consulted with DIBELS Data System (DDS) customer service at the University of Oregon to gain an understanding of which aspects of DIBELS first-hand users reported

©2026 University of Oregon. All rights reserved.

as the most valuable and the least valuable, as well as what they frequently ask for that DIBELS did not yet offer. As a result, several new goals were identified for DIBELS 8.

Increase the utility of NWF by expanding the spelling patterns assessed and grades in which it is available. Research (e.g., January, Ardoin, Christ, Eckert, & White, 2016) has demonstrated that NWF can be a more useful tool for screening and monitoring progress when patterns assessed move beyond CVC words. Research has also shown that this utility extends beyond kindergarten and first grade. As a result, expanding both the spelling patterns assessed by NWF and grades in which NWF was assessed became a goal.

Pay attention to order effects. Research (e.g., Burns et al., 2009) has shown the order in which items appear on fluency measures affects reading rate. Specifically, a form that begins with easier items and in which items become progressively more difficult allows for maximal performance by students, essentially by allowing them to gain momentum. In contrast, when item difficulty is more randomly distributed, the rate is adversely affected. As a result, using progressive difficulty as a principle in item assignments to forms became a goal for LNF, PSF, NWF, and WRF.

Add a word reading fluency measure. Research (e.g., Fuchs et al., 2004; January et al., 2016; Smith et al., 2014; Wise Justin et al., 2010) has also shown that NWF and ORF do not capture all struggling readers. Particularly in the early elementary grades, WRF improves prediction of students at risk, as well as monitoring of progress. As a result, the incorporation of WRF became a priority.

Provide consistent subtests within a grade. DDS customers frequently expressed a desire for consistency in subtests available within a grade. The lack of consistency, especially in kindergarten through second grade, made tracking the progress of all students during an academic year more challenging.

Validate specifically for use as a dyslexia screening tool. Perhaps the number one question of DDS customers in the last few years has been whether DIBELS is a valid screening assessment for dyslexia. While DIBELS has always been validated as a screener of risk for reading difficulties, it had never been specifically validated as a screening measure for word reading disabilities, including dyslexia. Of particular concern was the use of LNF as a measure of processing speed via rapid automatized naming (RAN). As a result, this new use of LNF, and DIBELS in general, informed development decisions as

©2026 University of Oregon. All rights reserved.

never before.

Letter Naming Fluency development process. The item pool for LNF consists of the uppercase and lowercase versions of all English letters, with the exception of the lowercase L (l) and both uppercase and lowercase W. Although these are obviously important letters for students to know, they were excluded to better align LNF to its increasing use as a RAN measure of processing speed. For the same reason, the kindergarten item pool was further limited to the 40 most frequently seen uppercase and lowercase letters, while the first grade pool includes the remaining nine letters.

To better control for differences in difficulty between forms, consistent rules are used in both kindergarten and first grade regarding when less frequent letters can appear on the forms. Each form in both grades begins with a sampling of the 20 most frequently seen letters, thereby preventing students from getting frustrated by forms that begin with rarer letters. Uppercase and lowercase letter frequency was determined based on the average frequency from five large corpora, as reported in Jones & Mewhort (2004). The 49 letters in the item pool were then grouped by average frequency into ten bins of five items each (with the exception of the final group, which includes only four items). In kindergarten, three sets of the top 20 items and two sets of items 21-40 were combined to create a total item pool of 100 letters per form. In Grade 1, two sets of the 49-item pool, plus two additional, randomly selected letters were combined to create a total item pool of 100 letters per form.

Each item was then assigned a random number using the default random number generator available in the statistical programming language R (R Core Team, 2018). Next, items were sorted so that each row included one letter from each bin, with the relative position of the individual letters in each bin determined by the random numbers assigned to each letter. Letters with a lower random number appeared before letters with a higher random number. Within each row, letters were strategically positioned so that the first row presented bins in decreasing order of frequency, and subsequent rows ordered the bins in varying combinations of difficulty.

This process, including the generation of a new set of random numbers, was repeated 50 times per grade to generate a pool of 50 potential forms. Multiple research staff reviewed each form, from which three were selected as the benchmark forms eliminating forms in which the same letter occurred more than once in succession or in which sequences of letters spelled English words.

©2026 University of Oregon. All rights reserved.

Phonemic Segmentation Fluency development process. To minimize the effect of vocabulary familiarity, all forms draw from the 2,500 most frequent two- to six-phoneme words in English, based on data from the English Lexicon Project (Balota et al., 2007). The initial item pool included all words from the English Lexicon Project that a) were identified as one of 2,500 most frequent words in both the Kučera & Francis (1967) word frequency list and the Hyperspace Analogue to Language (HAL) frequency norms (Lund & Burgess, 1996); b) had at least one meaning known by at least 50% of second grade students (Dale & O'Rourke, 1981); and c) had an adult-rated age of acquisition less than or equal to 7 (Brysbaert & Biemiller, 2017). This resulted in an initial pool of 662 words. We then removed homonyms (e.g., two, hear), potentially sensitive words (e.g., fight, hit), and two-phoneme words that were not among the 200 most frequent words. This resulted in a first-grade item pool of 594 words. For the kindergarten item pool, we further removed four- through six-phoneme words, resulting in a kindergarten item pool of 295 words.

To better control differences in difficulty between forms, consistent rules were used in both grades regarding where less frequent words can appear on the forms. Moreover, spelling patterns were ordered in terms of the number of phonemes, proceeding from two-phoneme words to words with progressively more phonemes. In kindergarten, the first six items have two phonemes, while the remaining 24 have three phonemes. In this way, PSF now avoids the distinct floor effects (i.e., many students scoring zero) in kindergarten that have plagued previous versions and, thus, eliminates the need for a separate measure of initial sound fluency. In first grade, the progression in difficulty is more rapid, with the first four items having two phonemes, the second six having three phonemes, and each subsequent group of six words having one more phoneme than the previous group.

All words in the final PSF item pool were assigned a random number using the default random number generator available in the statistical programming language R (R Core Team, 2018). This number was used to select words for inclusion in each section of the form (e.g., in kindergarten, the six two-phoneme words with the lowest random numbers were selected first, followed by the 24 three-phoneme words with the lowest random numbers). This process, including the generation of a new set of random numbers, was repeated 25 times per grade to create a pool of 25 potential forms. Multiple research staff reviewed each form, from which three were selected as the benchmark forms, and an

©2026 University of Oregon. All rights reserved.

additional 20 were selected as the progress monitoring forms.

Nonsense Word Fluency development process. The NWF item pool for DIBELS 8th Edition differs from previous versions of NWF in two important respects. First, all items now respect the English order and word position rules of individual letter combinations (Jones & Mewhort, 2004; Norvig, 2012), meaning that only phonetically regular letter combinations that actually appear in English are used: students are no longer asked to decode nonsense words such as fev or kaj. Second, the spelling patterns assessed have expanded beyond just the consonant-vowel-consonant (CVC) and vowel-consonant (VC) patterns used in previous versions. Although kindergarten forms are limited to CVC patterns, first grade forms include additional spelling patterns (described below) that are typically taught in first grade. DIBELS 8th Edition also now offers NWF in second and third grades, and includes additional, more complex phonics patterns in these grades.

The NWF item pool was created by first compiling lists of legal word parts in English, including various onsets and rimes. A total of 78 onset patterns were identified, including single letter onsets (e.g., b, s), blends (e.g., bl, tr), digraphs (e.g., ch, kn), trigraphs (e.g., str, thr), and VC onsets (e.g., am, ev). An additional 219 rimes were identified, including VC rimes (e.g., ab, in), vowel-consonant-e (VCe) rimes (e.g., abe, ide), vowel-r (Vr) rimes (e.g., ar, ir), vowel-r-consonant (VrC) rimes (e.g., arm, ort), vowel-consonant-consonant (VCC) rimes (e.g., est, olk), and vowel-vowel-consonant (VVC) rimes (e.g., aid, eed). These lists were then cross-combined in all possible legal English combinations and matched to a list of 31,845 real words and a separate list of 704 nonsense words that are pronounced like either a real word or a proper name or were deemed inappropriate or difficult to pronounce. Items on either list were removed from the item pool.

Finally, we used frequency counts of English letter n-grams (Norvig, 2012) to weight the relative frequency with which various word parts should appear on each form, computed a total frequency estimate for each nonsense word, and divided the total pool of nonsense words into quartiles based on that frequency estimate. Words with the least frequently appearing combinations of letters (i.e., those in the bottom quartile) were dropped from the pool, resulting in a total NWF item pool of 79,314 nonsense words.

©2026 University of Oregon. All rights reserved.

Items are arranged in five columns, and forms include between 75 (in kindergarten) and 100 (Grade 3) nonsense words. Form templates were created for each grade, which specified a particular sequence and relative frequency of each spelling pattern. In kindergarten, all 75 items are CVC words. In the first-grade template, the first five rows (a total of 25 items) consist entirely of CVC and VC nonsense words. In the next four rows, half of the items are again CVC or VC nonsense words, and the other half are silent-e (CVCe) and r-controlled (CVrC) nonsense words. In the next three rows, consonant blends (CVCC) and digraphs (CCVC) are introduced: 1/3 of items (i.e., 5 words) are VC and CVC, 1/3 are CVCe and CVrC, and 1/3 are CVCC and CCVC. In the final three lines, more complex patterns (i.e., CCVCC and CCCVC) are introduced: 1/3 of items are CVCe and CVrC, 1/3 are CVCC and CCVC, and 1/3 are CCVCC and CCCVC. The templates for second and third grades followed similar patterns but introduced additional spelling patterns: vowel digraphs (Grade 2), short vowel words ending in Y (Grade 2), and two-syllable words (Grade 3).

As with the other subtests, all nonsense words in the final NWF item pool were assigned a random number using the default random number generator available in the statistical programming language R (R Core Team, 2018). This number was used to select words for inclusion in each section of the form (e.g., in Grade 1, the 15 CVC and VC nonsense words with the lowest random numbers were selected first, and then the next 10 CVC and VC nonsense words were combined with the 10 CVCe and CVrC words with the lowest random number to complete the second section). Within each section, items were further randomized to ensure each word type appeared in a variety of positions. This process was repeated for each section, and then 36 times per grade (including the generation of a new set of random numbers) to create a pool of 36 potential forms. Multiple research staff reviewed each form, from which three were selected as the benchmark forms, and an additional 20 were selected as the progress monitoring forms.

Word Reading Fluency development process. Four different word lists contributed to the definition of the item pool for WRF. The Dale and O'Rourke (1981) word list is the only known list of words with age of acquisition determined by actual assessment with children. Because they assessed only fourth grade students and above, we supplemented their list with the far more recent work by Brysbaert and Biemiller (2017). Importantly, Biemiller estimated in earlier work (2010), that words

©2026 University of Oregon. All rights reserved.

known by 80% or more of fourth graders were likely to be known by 50% or more of second graders, thereby allowing for extension of the Dale and O'Rourke grade of typical acquisition down to second grade. These researchers asked adults to retrospectively estimate the age at which they knew words on the Dale and O'Rourke list. Results showed remarkable agreement between the children (test-based) and adults (retrospective) in terms of age of acquisition. As a result, we used the more fine-grained information from the newer list to further winnow down the list used for each grade.

Two word frequency lists were also used in creating the WRF pool. The Kučera and Francis (1967) word list is widely used for its comprehensiveness and availability. Nonetheless, it is an older list and the English language changes constantly. Thus, we also used the newer Hyperspace Analogue to Language (HAL) frequency norms (Lund & Burgess, 1996), which includes internet-based texts and is used very commonly as well. We consulted these word frequency lists together with age of vocabulary acquisition lists because most word frequency corpora, including the two we used, do not restrict themselves to children's texts. The combination of all four lists ensures that the words chosen have been frequent over several decades and are age- and grade-appropriate.

To create the WRF item pool, we began with a list of about 40,000 English words with a test-based age of acquisition rating (Dale & O'Rourke, 1981). We then excluded items that met any of four criteria words whose meaning was known by less than 50% of fourth grade students (Dale & O'Rourke, 1981). Next, we eliminated words with an adult-rated age of acquisition greater than 9 (Brysbaert & Biemiller, 2017). From this more limited pool, we further narrowed down the list to words with a frequency rating in both the Kučera & Francis (1967) word frequency list and the Hyperspace Analogue to Language (HAL) frequency norms (Lund & Burgess, 1996) that was greater than 7,500 per million words. Finally, words with at least one meaning identified as potentially inappropriate or distracting for students were omitted from the list. This resulted in a total item pool of 2,065 words for Grade 3.

Additional restrictions were imposed on the item pools for kindergarten through second grade. For all three grades, words had to be known by at least 50% of second grade students (Brysbaert & Biemiller, 2017). In Grade 2, words also had to have a frequency rating in the top 5,000 in both the Kučera & Francis (1967) and Lund & Burgess (1996) frequency norms, have an adult-rated age of acquisition less than or equal to 8, and could only be up to three syllables in length. The total item pool

©2026 University of Oregon. All rights reserved.

for Grade 2 was 1,111 words. In Grade 1, these criteria were further constrained. Namely, words had to have a frequency rating in the top 2,500 in both the Kučera & Francis (1967) and Lund & Burgess (1996) frequency norms, have an adult-rated age of acquisition less than or equal to 7, and be one or two syllables in length. The total item pool for Grade 1 was 652 words.

Finally, in kindergarten, the words had to have a frequency rating in the top 1,000 in both the Kučera & Francis (1967) and Lund & Burgess (1996) frequency norms, have an adult-rated age of acquisition less than or equal to 6 (Brysbaert & Biemiller, 2017), and could only be one syllable in length. The total item pool for kindergarten was 242 words.

In each grade, the item pool was grouped into three bins based on relative frequency. In kindergarten, words rated as one of the 50 most frequent words by both Kučera & Francis (1967) and Lund & Burgess (1996) were placed in the first bin, words with a frequency rating between 51 and 300 were placed in the second bin, and words with a frequency rating between 301 and 1,000 were placed in the third bin. In Grade 1, words rated as one of the 50 most frequent words by both Kučera & Francis (1967) and Lund & Burgess (1996) were placed in the first bin, words with a frequency rating between 51 and 1,000 were placed in the second bin, and words with a frequency rating between 1,001 and 2,500 were placed in the third bin.

In Grade 2, words rated as one of the 300 most frequent words by both Kučera & Francis (1967) and Lund & Burgess (1996) were placed in the first bin, words with a frequency rating between 301 and 2,500 were placed in the second bin, and words with a frequency rating between 2,501 and 5,000 were placed in the third bin. In Grade 3, words rated as one of the 1,000 most frequent words by both Kučera & Francis (1967) and Lund & Burgess (1996) were placed in the first bin, words with a frequency rating between 1,001 and 5,000 were placed in the second bin, and words with a frequency rating between 5,001 and 7,500 were placed in the third bin.

Items were then assigned a random number using the default random number generator available in the statistical programming language R (R Core Team, 2018) and arranged by frequency bin and random number. In kindergarten, the 15 words in the first frequency bin with the lowest random numbers were selected as the top three rows of the form, the 35 words in the second frequency bin with the lowest random numbers were selected as the next seven rows, and the 35 words in the third

©2026 University of Oregon. All rights reserved.

frequency bin with the lowest random numbers were selected as the last seven rows. In Grade 1, the 15 words in the first frequency bin with the lowest random numbers were selected as the top three rows of the form, the 45 words in the second frequency bin with the lowest random numbers were selected as the next nine rows, and the 45 words in the third frequency bin with the lowest random numbers were selected as the last nine rows.

In Grade 2, the 20 words in the first frequency bin with the lowest random numbers were selected as the top four rows of the form, the 55 words in the second frequency bin with the lowest random numbers were selected as the next 11 rows, and the 55 words in the third frequency bin with the lowest random numbers were selected as the last 11 rows. In Grade 3, the 30 words in the first frequency bin with the lowest random numbers were selected as the top six rows of the form, the 55 words in the second frequency bin with the lowest random numbers were selected as the next 11 rows, and the 55 words in the third frequency bin with the lowest random numbers were selected as the last 11 rows.

This process, including the generation of a new set of random numbers, was repeated 30 (in kindergarten and Grade 1) to 40 (in Grades 2 and 3) times per grade to create a pool of potential forms. Multiple research staff reviewed each form, from which three were selected as the benchmark forms, and an additional 20 were selected as the progress monitoring forms.

Oral Reading Fluency development process. Rather than hiring item writers to author the new ORF passages, we hired published and aspiring short story authors: Rose Gowen, Kristen Havens, Sarah Meacham, Ben Seipel, Bob Thurber, Tina Truitt, and Andrew Wilson. Rose Gowen is an American writer and mother of two living in Montreal who has been published in the American Poetry Review, Night Train, and McSweeney's among other venues and attended the 2018 Bread Loaf Writers' Conference. Kristen Havens is a writer and editor living in Los Angeles, CA, who has written for many clients including the Special Olympics, has received multiple honorable mentions from Glimmer Train, and was nominated for the PEN/Robert J. Dau Short Story Prize for Emerging Writers. Sarah Meacham is a writer, anthropologist, and mother living in Los Angeles, CA, who was a staff writer for the UCLA Division of Social Sciences and External Affairs and worked with the Strategic Education Research Partnership in Boston Public Schools. Ben Seipel is an Assistant Professor at California State University, Chico,

©2026 University of Oregon. All rights reserved.

and is an aspiring author who taught Spanish in K-12 in Minnesota for many years. Bob Thurber is an author and father living in North Attleboro, MA, who has published two novels and innumerable short stories, appeared in over 50 short story anthologies, and won more than 20 writing awards. Tina Truitt is an author, mother of three, and preschool teacher living in Cherry Hill, NJ, who has published two books, including a children's multicultural, bilingual picture book about teamwork. Andrew L. Wilson is an author and editor living in Eugene, OR, who has published poetry and short stories in a wide range of venues, including *Exquisite Corpse* and *In Posse Review*, and has edited the online literary journal Linnaean Street as well as academic books and technical reports. The authors come from diverse socio-economic and cultural backgrounds.

Authors were given detailed specifications to guide them in writing their assigned passages, which included narrative and informational texts for multiple grade levels. Specifications for passage length and Flesch-Kincaid grade-level readability were also provided (see Table 1.6 on page 31). Authors were coached to represent diverse experiences in terms of culture, geography, and locale, as well as to avoid hackneyed and culturally sensitive topics.

In addition, authors were asked to give each passage a relatively short title that did not give away the ending, as well as use standard English formatting and grammar and grade-level appropriate topics and vocabulary. Narrative texts were required to have a discrete beginning, middle, and end, with multiple episodes or events in the middle. Informational texts were required to have a clear introduction and conclusion with intermediate paragraphs that provided supporting details, and where possible utilize text structures frequently used in the elementary grades (i.e., compare-contrast, cause-effect, problem-solution, and sequence). Authors were also asked to avoid dialogue, headings, slang, italics, and bold font, as well as content that could be considered religious, controversial, or offensive to some cultures. Finally, they were encouraged to refrain from writing passages that were too funny or emotional, consistently similar in style and tone, or overly arcane or familiar in topic.

©2026 University of Oregon. All rights reserved.

Table 1.6 DIBELS 8th Edition Oral Reading Fluency Passage Writing Specifications

Grade	Required length in words	Target Flesch-Kincaid grade level
1	150-225	1.5-2.0
2	150-225	2.5-3.0
3	175-275	3.5-4.0
4	175-275	4.5-5.0
5	200-325	5.5-6.0
6	200-325	6.5-7.0
7	250-375	7.5-8.0
8	250-375	8.5-9.0

Once passages were turned in by the authors, the DIBELS 8th Edition development team reviewed them for consistency with the specifications. In cases where passages diverged from these specifications, passages were revised by the DIBELS 8th team in cases where the passage was deemed salvageable. Others were discarded at this stage. Grade level was determined by readability level (i.e., Flesch-Kincaid grade level).

Next, all passages were reviewed by a team of external reviewers who were parents and/or former teachers with experience with K-8 students and settings. Reviewers were trained by familiarizing them with oral reading fluency measures and the purposes of the review, as well as the criteria by which they would evaluate stories. They reviewed the passages for grade-level appropriateness of their vocabulary, syntax, sentence length, and overall content, as well as the background knowledge required for comprehension.

They also indicated when passages were likely to evoke an emotional reaction from readers that might interfere with reading rate (e.g., laughing out loud, gasping in surprise). In addition, they were

©2026 University of Oregon. All rights reserved.

asked to rate passages for how accessible and enjoyable they were for slow and struggling readers, helping to ensure that the first few sentences were not overly difficult and provided a hook (or schema) that supported comprehension. Furthermore, they reviewed passages for potential bias, indicating whenever they judged a passage as potentially offensive to readers or teachers based on gender, ethnicity, race, national origin, religion, disability status, sexual orientation, and geographical region. They were also asked to rate potential for bias due to passage topic and tone, especially bias toward students from backgrounds typically under-represented in children's texts. Note that potential bias in ORF passages was also addressed through sensitivity analyses of classification accuracy for readers for different backgrounds. This information can be found in the DIBELS 8 Technical Manual.

Finally, reviewers indicated if a given passage might be as or more appropriate for other grade levels. Importantly, the training emphasized that reliability of ratings was not a goal and diversity of opinions was perfectly acceptable.

Once passages had been reviewed by two or more of the panel members, DIBELS 8th Edition researchers analyzed ratings and revisited all passages where reviewers noted one or more problems. In some cases, passages were immediately discarded. Judgments regarding vocabulary inappropriateness were supplemented with checks of word frequencies and age of acquisition, and in cases where the inappropriateness was confirmed, a more grade-appropriate substitution was made. Judgments regarding syntactic complexity resulted almost uniformly in similar revisions.

Of particular importance was the content appropriateness, which resulted in passages being considered for assignment to higher and lower grade levels than their readability would suggest. These judgments were sometimes based on the background knowledge required to comprehend a passage, but also often relied on the sophistication of literary and rhetorical devices and overall conceptual complexity. In such cases, some effort was made to increase or decrease readability to improve apparent "fit" with the new grade level assignment. Nonetheless, current consensus is that the appropriate grade level of reading material is more than a matter of strict readability. Thus, given that oral reading fluency is intended to act as an indicator of reading comprehension (rather than strictly of decoding skill efficiency), some passages were assigned to higher and lower grade levels even when readability did not strictly match the assigned grade (see Appendix A on page 107).

©2026 University of Oregon. All rights reserved.

Finally, all passages were field-tested in their targeted grade levels. Passages where reviewers disagreed about text complexity and grade appropriateness were field-tested in multiple grades. The final assignment of passages to grades and benchmark periods was based on student performance on the passages, the predictive validity of specific passages in a given grade, and maintaining a balance of narrative and informational texts. We increased the diversity of narrative and informational subgenres represented across the intermediate and middle grades. Among the subgenres we included in these grades are fantasy, science fiction, western, and mystery passages. For informational texts, we increasingly varied topics across content areas (e.g., life sciences, earth sciences, ancient history, modern history, biography) and also varied text structures (e.g., compare and contrast, cause and effect, description, problem and solution, and procedural). In many cases, it is difficult to categorize a passage as narrative or informational; for instance, a passage written in the first person about an informational topic could be considered narrative, informational, or an amalgam of both depending on its particular style. Nonetheless, in Grades 1-5, we required that narrative passages make up more than half the passages with an average balance of 60% narrative to 40% informational. In Grades 6-8, we relaxed this requirement and selected more informational passages with an average balance of 40% narrative to 60% informational. Key text readability and complexity statistics are reported for all benchmark passages in Appendix A on page 107.

Maze development process. Maze passages were developed in the same manner as ORF passages but went through a few additional steps of development. First, passages were lengthened to reach typical lengths found in other CBMs and in previous DIBELS editions to allow for enough items for appropriate measurement of readers with better fluency and comprehension.

Second, following common rules, the first and last sentences of every passage were left intact, except in Grade 2 where the second sentence was also left intact to allow for better establishment of a situation model for the passage (Kintsch, 1998). Third, beginning with the third word of the second sentence (or third sentence in Grade 2), every seventh word was deleted with a few caveats. If the seventh word was a proper noun or number, then the eighth word was deleted. If the seventh word was highly specialized (e.g., an uncommon scientific term for a given grade), it would not be deleted unless it had occurred previously in the passage. Also, hyphenated words were treated as one word.

©2026 University of Oregon. All rights reserved.

Third, the deleted word became one of the answer choices, and two distractors were written for each deleted word. Each distractor was written by a different DIBELS 8th Edition researcher according to a number of rules informed by research. Distractors could not begin with the same letter as the correct word (Conoyer et al., 2017). Distractors were also kept to within two letters in length of the correct answer, although this rule was relaxed in the upper grades (i.e., Grade 5 and beyond). When the deleted word was a noun, verb, or adjective, distractors had to be grammatically correct. For instance, if the word to be chosen followed "an", then the distractors had to begin with a vowel. When the deleted word was a contraction, all distractors also had to be contractions and tense agreement was deemed unimportant. Different forms of the same word were never used as distractors (e.g., "be", "is", and "are"). For all other parts of speech, grammatical correctness was not a requirement because it was found to result in repetitive distractors. For example, when the deleted word was an article, requiring grammatical correctness resulted in the answer choices always being "a", "an", and "the." It was deemed undesirable to have answer choices repeat too frequently. Finally, in Grade 5 and up, one of the distractors was required to have semantic similarity to the correct word. That is, it could make sense in a given sentence but not in the story as a whole.

Once distractors were written, they were reviewed by another DIBELS 8th Edition researcher, who would make corrections when rules were violated. If the reviewer found a particular item to be inordinately difficult, the item was brought to a subset of researchers for discussion and potential revision. Finally, the answer choices were reordered so that they were always listed alphabetically.

Benchmark passages were selected from the resulting pool using rules that balanced readability, text complexity, and Lexile ratings (see Table 1.7). In order to balance these factors, readability grade levels were permitted to go above grade level in all but second grade. Key text readability and complexity statistics are reported for all benchmark passages in Appendix B on page 115.

©2026 University of Oregon. All rights reserved.

Table 1.7 DIBELS 8th Edition Maze Benchmark Passage Selection Specifications

Grade	Required length in words	Target Flesch-Kincaid grade level	Lexile	Coh-Metrix narrativity score
2	350+	2.0-2.9	500L – 600L	80+
3	350+	3.0-4.9	500L – 600L	70+
4	400+	4.0-5.9	700L – 900L	60–90
5	400+	5.0-7.5	800L – 1000L	50–80
6	400+	6.0-8.5	900L – 1100L	20–70
7	450+	7.0+	900L – 1100L	20–70
8	450+	8.0+	1000L – 1200L	< 70

Summary

This chapter has laid out the history and most recent developments of DIBELS measures. As noted throughout, DIBELS researchers based decisions about DIBELS 8th Edition on the research literature, user feedback, and ongoing research conducted by the University of Oregon (UO). Research into the properties of DIBELS and how to improve its usefulness is ongoing at UO. Regular addendums to this manual will keep DIBELS 8 users up-to-date on the features and technical qualities of DIBELS.

©2026 University of Oregon. All rights reserved.

Chapter 2: Administration Instructions and Scoring Procedures

DIBELS® 8th Edition is intended for use with students enrolled in kindergarten through eighth grade. Subtests can be administered to students with or without reading difficulties and disabilities, with frequency of assessment adjusted based on the assessment purpose (e.g., universal screening, progress monitoring).

Examiners who give and interpret DIBELS 8th Edition must receive training in standardized administration and scoring procedures. Standardization ensures reliable scores and allows for comparisons between results and research-determined criteria.

The next section presents general guidelines for administering DIBELS. That section is followed by specific instructions for administering and scoring the six DIBELS 8th Edition subtests: (a) letter naming fluency, (b) phonemic segmentation fluency, (c) nonsense word fluency, (d) word reading fluency, (e) oral reading fluency, and (f) maze. Specific materials required for each subtest are listed in the subtest descriptions. Throughout this chapter, bold font is used to indicate scripted directions or prompts provided to the student.

DIBELS 8th Edition General Guidelines

There are a number of common features across the DIBELS 8th Edition individually-administered subtests. For each subtest, the following are provided:

Applicable grades: the grades for which a subtest is designed;

Objective: the activity in which the student engages, including administration time;

Uses: the uses for which the subtest is designed.

In addition, a list of required materials is provided for each subtest, along with detailed administration instructions. Instructions include when to start and stop timing, how to score, and the allowed reminders and prompts. Numerous scoring examples are also given for each subtest.

All DIBELS subtests are best administered in a quiet location where minimal interruptions can be expected. For individually administered measures (i.e., all but Maze), a table or desk separated from distractions is best. When individually administered measures are conducted in a classroom setting,

©2026 University of Oregon. All rights reserved.

other students should be engaged in quiet, independent activities. The assessment administrator should sit so that the student's face is easily seen and close enough to easily point to forms and hear what the student says. No matter how close the student and administrator sit, the scoring booklet should not be visible to the student, which is why we recommend using a clipboard.

General Timing and Scoring Guidelines

DIBELS is a timed measurement system. With the exception of Maze, all DIBELS 8th Edition subtests are 60-second timed measures. Maze is a 3-minute timed measure. All DIBELS measures are timed because DIBELS is designed to be a measure of fluency. In all cases, it is critical to time each administration as accurately as possible. Even small mistakes can result in less reliable, and thus less valid scores, and research has shown that timing mistakes are among the most common (Reed, Cummings, Schaper, Lynn, & Biancarosa, 2018). For DIBELS to be a valid assessment, strict adherence to timing conventions is required.

Scoring for all the 60-second subtests has certain commonalities. When 60 seconds have elapsed, the examiner always places a bracket (i.e.,]) after the last item completed and says, "Stop." Also, if a student makes an error, put a slash (i.e., /) through the incorrect item. If a student makes an error but self-corrects the error within 3 seconds, mark SC over the item.

For all subtests with student materials, if the student is responding in a way that the examiner loses track of which stimuli the student is responding to, it's acceptable for the examiner to quickly redirect the student by pointing. This requires professional judgment and the examiner should consider whether the administration should be invalidated. (See the Invalidating Administrations section on p. 43.)

Order of Administration

In kindergarten and the beginning of first grade, we strongly recommend administering the subtests in the order of skill development. Begin with LNF, which should be followed by PSF, then NWF, then WRF, and in first grade then ORF. An illustration of the suggested order of administration, from left to right, with discontinue rules is below.

©2026 University of Oregon. All rights reserved.

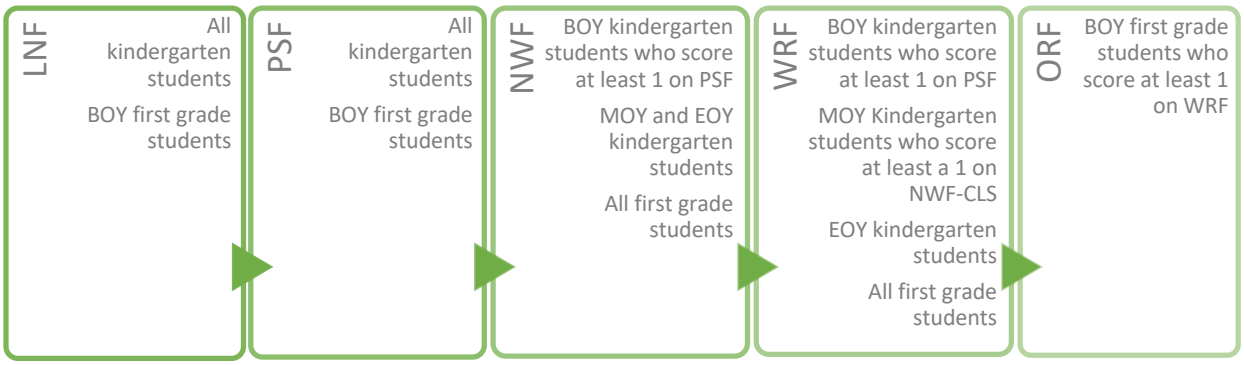


Figure 2.1 Order of test administration from kindergarten through the beginning of first grade.

From the middle of first grade on, subtests assessing lower-level skills should only be administered after the gating subtest is given. In first grade, the gating subtest is NWF, and in second and third grade the gating subtest is ORF. Suggested order of administration, from right to left, and gating rules are summarized in the figure below.

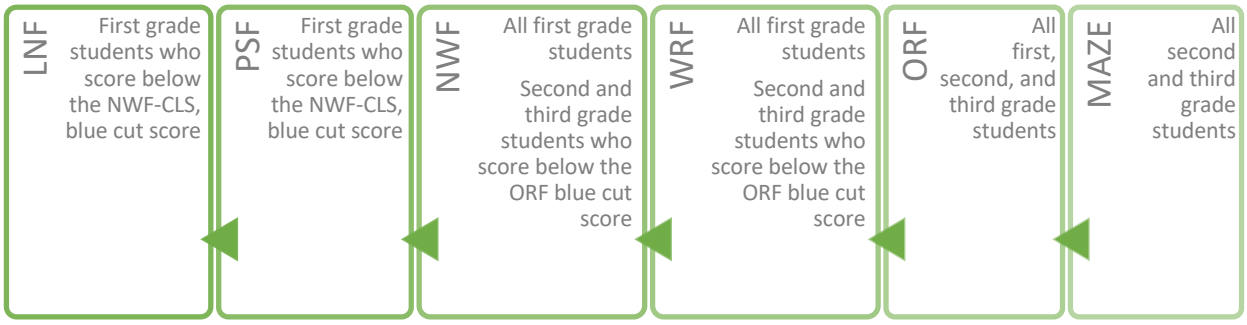


Figure 2.2 Order of test administration from middle of first grade through third grade.

Benchmarking Discontinue and Gating Rules

Each DIBELS subtest has a specific discontinue rule. An assessment should only be discontinued if the specified conditions have been met, or if the administration is irrevocably interrupted (e.g., a fire drill occurs). See the rules for each subtest for its discontinue criteria.

For some subtests at certain time points, not only is that subtest discontinued, but benchmark assessment is also discontinued altogether. DIBELS 8th Edition offers discontinue benchmarking rules for kindergarten and first grade. These rules prevent unnecessary and excessive testing for the most vulnerable learners by giving educators the option to stop the administration of benchmark assessments based on student performance. For example, in the beginning of kindergarten, we

©2026 University of Oregon. All rights reserved.

recommend that testing stop if a student is unable to segment words phonemically. Our research has shown that students who score zero on PSF are only extremely rarely able to score any points on NWF or WRF, making the administration of these subtests highly unlikely to yield additional useful information. The rules were derived from a national field trial that indicated students who scored 0 for the indicated assessments in the periods specified above were extremely unlikely to get any items correct on the remaining subtests. Nonetheless, examiners have the option of administering the remaining subtests based on professional judgment. The benchmark discontinue rules for Kindergarten and beginning of First Grade are explained where applicable and are also summarized in Table 2.1. It is important to note that there is only one subtest used for the Discontinue Benchmarking Rule in each benchmark period, but discontinuation of individual subtests can still happen. The Discontinue Benchmarking Rule is to discontinue all benchmark administration. For example, if a student scores a one on PSF at the beginning of kindergarten, then you would administer both NWF and WRF, regardless of the student's performance on NWF. Thus, even if the student gave no correct answers in the first five NWF words, you would still administer WRF.

Table 2.1 Benchmark Discontinue Rules

Grade and Time of the Year	Benchmark Discontinue Rule	Scoring
Kindergarten, Beginning	If PSF is discontinued, do not administer NWF and WRF.	Enter 0 for PSF. Do not enter scores for the remaining subtests: NWF and WRF
Kindergarten, Middle	If NWF is discontinued, do not administer WRF.	Enter 0 for NWF. Do not enter scores for the remaining subtest: WRF
First grade, Beginning	If WRF is discontinued, do not administer ORF.	Enter 0 for WRF. Do not enter scores for the remaining subtest: ORF

©2026 University of Oregon. All rights reserved.

We have introduced additional gating rules that are also designed to prevent unnecessary and excessive testing for all learners. Beginning in the middle of first grade and extending through the end of third grade, we recommend that students who are at negligible risk (i.e., score above the ambitious cut) based on a specific subtest not be tested with subtests tapping lower level skills. In first grade, students who score at or above the ambitious cut on NWF-CLS need not be given LNF or PSF at the middle or end of the year. In second and third grade, students who score at or above the ambitious cut on ORF-WRC need not be given NWF and WRF. We do not offer gating rules beyond third grade, but we will continue to investigate ways to introduce testing efficiencies in these grades. Finally, as with the discontinue benchmarking rules, educators always have the option to administer subtests despite a student qualifying for gating. The benchmark gating rules are explained where applicable and are also summarized in Table 2.2.

Table 2.2 Benchmark Gating Rules

Grade and Time of the Year	Benchmark Gating Rule	Scoring
First grade, Middle and End of Year	If NWF is above the blue cut, do not administer PSF or LNF.	DDS and mCLASS users: Do not enter scores for PSF and LNF. Others: Enter the last green score (the blue cut-score minus one) for PSF and LNF.
Second and Third grades, All year	If ORF is above the blue cut, do not administer NWF or WRF.	DDS and mCLASS users: Do not enter scores for NWF and WRF. Others: Enter the last green score (the blue cut-score minus one) for NWF and WRF.

The discontinue and gating rules have two important ramifications for the administration and scoring of DIBELS 8. The first relates to the order of subtest administration, and the second to the computation of composite scores for students who are discontinued or gated.

©2026 University of Oregon. All rights reserved.

While the gating rules are designed to save on unnecessary testing time, their use is optional. If you want to track growth on a specific measure, then that measure should be administered at all time periods regardless of the gating rules. Likewise, if you want to track growth on the composite score, then you should carefully consider whether to utilize the gating rules. Note that we used the end of the green (or benchmark) range of scores because there is a decent chance that a gated student may score at the benchmark level but not the ambitious level for the gated measures. Thus, while we provide substitute scores in Table 2.2 for students who are discontinued, when there is a need to track growth as precisely as possible, administering all measures will result in the most precise subtest and composite scores for tracking growth.

Invalidating Administrations

An important aspect of administering any assessment is knowing when an administration ought to be treated as invalid. For a fluency-based assessment like DIBELS 8, many things can occur that would ruin an administration. In such cases, a score should not be entered and an alternative form should be administered at another time. The challenge here is deciding when an administration has indeed become invalid and choosing an alternative form to administer.

When to Invalidate an Administration

Situations and errors that spoil an administration include, but are not limited to, the student refusing to participate, the student being too ill to participate, the administrator forgetting to start the timer or missing the end of the 60-second period, and situational interruptions.

One of the most common examples of such a situation is when a fire drill occurs in the middle of an administration. Because timing is central to DIBELS scoring, the distraction alone is enough to invalidate the administration. In other words, even if the alarm were turned back off within seconds, the student's (and test administrator's) attention has been irrevocably distracted.

It is important to be sensitive to less common situations that can also ruin an administration. A student may refuse to comply with instructions, such as when a student who can and does read in other contexts refuses to read aloud. A student may be overcome with emotion, such as when a student who is struggling inordinately with a task begins to cry. Maze also can involve some unique situations, including when a student skips a page by accident or receives a packet where not all pages

©2026 University of Oregon. All rights reserved.

are included.

In each of these cases, and more than we can list here, test administrators need to use their best professional judgment as to whether (a) an administration has, in fact, become invalid and (b) a new administration ought to be undertaken. For example, a student who refuses to read should be assessed another day, when perhaps the student will be more compliant. In the example of a student crying, the test administrator needs to decide whether the situation was the result of a bad day, in which case a new administration would be advisable and no score entered for the current administration. Alternatively, the administrators may decide crying was the result of a task simply being too difficult, in which case a new administration is not advisable and the achieved score should be entered. In the Maze cases described, an alternative form should always be administered.

Choosing an Alternative Form

When a new administration is necessary, best practice dictates using a progress monitoring form for the subtests for which these forms are available. At the beginning of the year, using the first progress monitoring form is easiest. When a student has already undergone progress monitoring to any extent with a subtest, choose a form that the student has not yet seen. If a student has seen all the progress monitoring forms, go back to the first progress monitoring form.

LNF creates a particular challenge because of the lack of progress monitoring forms. In this case, administering a benchmark form from a different time of year for the same grade level is advisable. If a week or more passes between the invalidated administration and the new one, the identical form can be used. Because LNF is not a meaning-laden task (in contrast to reading passages for ORF or Maze), experiencing an LNF form more than once is less of a problem, so long as sufficient time has passed for a student to forget what was seen before.

When administering any form that is not the intended form for a given benchmark time of year, it is critical to note the actual form used. That information will help avoid reusing forms during any later progress monitoring.

In all cases where an administration is invalidated, the student should be reassessed using a progress-monitoring form at another time or on another day depending on the administrator's professional judgment. In general, though, the student should be assessed as soon as possible.

©2026 University of Oregon. All rights reserved.

Giving Instructions and Encouragement

DIBELS 8th Edition is a standardized assessment, which means test administrators must adhere to scripted procedures for giving students directions in addition to following the timing rules. Test administrators should only say what is provided in the administration instructions and should speak clearly enough for the student to hear well.

Students should not be given feedback on their performance during or after an assessment. If an examiner wishes to give a student general encouragement in between subtests, praising the student's effort is acceptable (e.g., "Nice effort! I can see you're working to do your best.").

For many subtests, practice items are provided. Again, the test administrator should adhere to the scripted instructions. Offering additional practice, corrections, or off-script explanations is not allowed. Because DIBELS is an assessment and is used for instructional decision-making, it is critical to determine a student's performance without undue instruction or intervention. The practice items only serve the purpose of ensuring that students comprehend the task at hand.

If a student clearly does not hear or understand instructions or practice items, the test administrator may repeat these procedures once. If the assessment has already begun, the timer should be kept running.

Important Considerations for Inclusive and Accurate Scoring

A difficult aspect of scoring DIBELS and similar assessments that require students to respond orally is scoring accurately while maintaining inclusive practices that respect culturally and linguistically diverse students. This section highlights some of the most important considerations regarding respecting and not penalizing students for their accents, dialects, articulation, and instructional histories in scoring. The issues discussed here tend to affect PSF and NWF most directly, but can also influence scoring of WRF and ORF.

Phonemes, Phones, and Scoring PSF and NWF

DIBELS 8th Edition's phoneme pronunciation guide has some key differences from previous DIBELS phoneme pronunciation guides. We made these changes to have DIBELS scoring more accurately reflect the phonemic structure of English. These changes specifically address r-controlled vowels and diphthongs.

©2026 University of Oregon. All rights reserved.

We often speak of phonemes as the smallest unit of sound in a language, but it's actually a little more complicated than that. Phonemes are the smallest unit of sound in a language that distinguish one word from another (i.e., if one phoneme is swapped for another, it changes the meaning of the word). Many phonemes actually consist of more than one phone. Phones are ANY distinct sound in speech.

Both r-controlled vowels and diphthongs are single phoneme sounds in English that contain two phones. In other words, the natural perception in English is of one sound. Nonetheless, we exaggerate the phoneme into phones when learning to read and write.

DIBELS 8th Edition scoring rules reflect the natural phonemic perception in American English, meaning all of the words are split into their phonemes and not phones in the scoring guides. For example, farm is /f/ /ar/ /m/, and coin is /k/ /oy/ /n/ using the DIBELS 8 phoneme pronunciation guide. These new rules create a scoring conundrum in many cases. What do we do when students further segment r-controlled vowels and diphthongs accurately into their constituent phones?

When students segment r-controlled vowels and diphthongs accurately, they should be scored as correct because they are actually subdividing sounds below the phonemic level; that is, they are accurately segmenting the English phonemes into their constituent phones and should not be penalized for this practice. A few examples are offered below to illustrate how this works in practice.

Let's start with a long A diphthong using the word "ray." Ray is two phonemes in English: /r/ /A/ (using the DIBELS 8 phonemic notation). The /A/ actually includes two phones: the short e (/e/ in DIBELS 8 notation) and the long e (/E/ is DIBELS 8 notation). Phonologically, there really ARE two sounds in the long A. But in English we don't teach students about that because the distinction is meaningless in English (i.e., it literally has no impact on the meaning of ray). Note, too, that both /e/ and /E/ are phonemes in English, but they do not play that role when combined in English words. In essence, they "become" /A/. Native Spanish speakers typically hear these sounds and often break up /A/ into its constituent phones because Spanish does not include the long A sound as a phoneme. In contrast to what we'll see with r-controlled vowels, hearing /e/ /E/ is not useful in trying to spell in English, so we don't sensitize our students to the fact that there are two phones in the phoneme /A/.

©2026 University of Oregon. All rights reserved.

Where diphthongs can become confusing is when we do sensitize our students to them. For example, the vowel sound in “toy” is a single phoneme in English: /oy/ (using DIBELS 8 phonemic notation). Nonetheless, many curricula emphasize breaking the diphthong into its constituent phones, most commonly /O/ and /E/ (using DIBELS 8 notation), to support spelling.

R-controlled vowels in English work much the same way. We naturally hear r-controlled vowels as a single unit: are, or, air, ear, etc. As with ray, we can be trained to hear the individual phones that make up r-controlled vowels (i.e., quite literally separating the /r/ from the vowel sounds). Phonemically this distinction is again meaningless (it makes no difference in what the word means if we “hear” /air/ or /A/ /r/, the same as it makes no difference if we “hear” /ar/ or /ah/ /r/). In contrast to long vowel sounds in English, in the case of r-controlled vowels, it can be useful to sensitize students to the phones that make up these phonemes because it helps with spelling.

Previous editions of DIBELS used to count the r-controlled vowel sound in “are” as one phoneme (/ar/), but the r-controlled vowel sound in “air” as two phonemes (/A/ /r/). Unfortunately, this practice led to substantial confusion. Thus, we have adjusted our phonemic pronunciation guide to strictly and accurately represent the phonemic structure of English. As a result, with DIBELS 8th Edition we score for what matters phonemically in English, so that all r-controlled vowels are represented as one phoneme.

Many more examples exist in which students may generate more phones during NWF assessment than exist at a phonemic level. As a reminder, though, if a child segments a word accurately into phones (below the phoneme level, saying /A/ /r/ for /air/ for example), we do not penalize them! We consider them correct. Students should not lose points for finer-grained segmentation, as long as the correct phones are used. In the same way, a Spanish speaker who reliably segments /A/ into /e/ and /E/ should get credit as well.

Articulation, Accent, and Dialect

DIBELS 8th Edition measures early literacy skills in English. Therefore, students should use the English pronunciation of words. However, it is important to mention that students are not penalized for varied pronunciation due to consistent dialect, accent, or articulation differences. For example, if the student consistently says /th/ for /s/ and pronounces “thee” for “see” when naming the letter “C”, credit is given for naming the letter correctly. This is a professional judgment and should be based on

©2026 University of Oregon. All rights reserved.

the student's responses and any prior knowledge of the student's speech patterns.

Different regions of the country use different dialects of American English. The DIBELS 8th Edition Phoneme Pronunciation Guide (see Appendix C on page 122) is particularly helpful with the Phonemic Segmentation and Nonsense Word Fluency subtests. These pronunciation examples may be modified consistent with regional dialects and conventions.

An important update to the DIBELS pronunciation guide is the treatment of r-controlled vowels (e.g., word, far), which are sometimes also called r-colored vowels. Considerable disagreement exists about how many phonemes exist in words with r-controlled vowels and thus in American English (e.g., Bizzocchi, 2017; Fry, 2004; Lockenvitz, Kuecker, & Ball, 2015). Whereas earlier editions treated some as single phonemes and others as two or more phonemes, DIBELS 8th Edition simplifies the treatment of r-controlled vowels by treating them as single phonemes. Nonetheless, test administrators should take local dialects and articulation issues into account when scoring nonsense words or phonemic segmentations involving r-controlled vowels. In some regions in the US, r-controlled vowels are more clearly separated into multiple sounds or phonemes (e.g., "lair" might be pronounced as "layer"). Students using such a pronunciation should not lose points for this practice.

Accommodations

The DIBELS 8th Edition measures are designed to be used unmodified with all students. They have been validated with thousands of students using the DIBELS 8th Edition standardized procedures. Interpretation of student scores is only informative when students have been assessed in this standardized way.

In a very small number of cases, however, several **accommodations** are approved. These accommodations should only be used in situations where they are necessary to obtain an accurate score for a student. In other words, accommodations should only be used if there is evidence that without them, the assessment would be measuring something other than the intended reading-related skill. For example, if a student is hard of hearing and without an accommodation the student would not be able to hear the testing directions, then that would result in the test measuring the student's hearing abilities rather than reading skills. An accommodation would be appropriate in this case.

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition-approved assessment accommodations involve minor changes to assessment procedures that are unlikely to change the meaning of the results and have been approved either by DIBELS developers or assessment professionals. They should be used only when:

- An accurate score is unlikely to be obtained without the accommodation; and/or
- Specified in a student's 504 plan or Individualized Education Plan (IEP).

The accommodations approved for DIBELS 8th Edition are listed in Table 2.3. When approved accommodations are used, the examiner should mark an "A" on the front cover of the testing booklet. Scores from tests administered with accommodations can be compared to other DIBELS 8th Edition benchmark scores and norms. *Approved accommodations should only be used with students who have a documented need for such supports, such as a 504 plan or IEP, or knowledge that an accurate score could not be obtained without the accommodation (e.g. Spanish directions).*

Table 2.3 Acceptable Accommodations for DIBELS 8th Edition

Accommodation	LNF	PSF	NWF	WRF	ORF	Maze
Quiet setting for testing	X	X	X	X	X	X
Breaks in between measures	X	X	X	X	X	X
Assistive technology (e.g., hearing aids, assistive listening devices, glasses)	X	X	X	X	X	X
Administration Directions in the student's home language (currently available in Spanish)	X	X	X	X	X	X
Enlarged student materials	X		X	X	X	X
Colored overlays, filters, or lighting adjustments	X		X	X	X	X
Marker or ruler for tracking	X		X	X	X	X
Whisper phones						X
Slant board	X		X	X	X	X

©2026 University of Oregon. All rights reserved.

Anything an assessor does that is not listed in the standardized scoring and administration and is not an approved accommodation falls under the category of a **modification**. Any modification made to the standardized directions, timing or scoring rules renders results that are likely to be meaningfully different than they would have been without the modification. Examples of unapproved accommodations and modifications include: (a) extending the time on a DIBELS probe, (b) repeating practice items, (c) providing different or extra models of the task, (d) adding to or changing administration directions, and (e) offering unapproved prompts and feedback.

When unapproved accommodations or modifications are used, the examiner should mark an "M" on the front cover of the testing booklet. Scores are *not* valid in these cases and should not be entered in a data system or interpreted in relation to DIBELS 8th Edition benchmark goals and norms.

It is important to recognize that there are some students for whom DIBELS is not an appropriate assessment. Students for whom this is true include those:

- With limited verbal language skills,
- With fluency-based speech disorders or oral apraxia, and/or
- For whom reading in English is not an instructional goal (e.g., students learning to read exclusively in a language other than English).

In these cases, other assessments and curricular tools (e.g., end-of-unit tests, individualized progress monitoring materials, other-language reading assessments) are best suited to screening students and monitoring student progress toward goals.

©2026 University of Oregon. All rights reserved.

Letter Naming Fluency (LNF)

Applicable grades: Beginning of kindergarten through end of first grade.

Objective: Student names letters for 60 seconds.

Uses: Benchmark and risk assessment.

Materials

- Scoring book
- Student form
- Pen or pencil
- Clipboard
- Timer

Administration

1. Position the clipboard and scoring book so that the student cannot see what you record.
2. Place the student copy of the LNF subtest in front of the student.
3. Say these specific directions:

   Here are some letters

   (point to the student form).

   Tell me the names of as many letters as you can.

   When I say "Begin," start here,

   (point to the first letter)

   and go across the page

   (point).

   Point to each letter and tell me the name of that letter.

   If you come to a letter you don't know, I'll tell it to you.

   Put your finger on the first letter.

   Ready?

   Begin.

4. Start the timer after saying "Begin."
5. Follow along in the Scoring Booklet. Put a slash (/) through each letter name read

©2026 University of Oregon. All rights reserved.

incorrectly. See Acceptable Prompts and Scoring Rules for more details.

6. At the end of 60 seconds, place a bracket (]) after the last letter named and say, "**Stop**."

Acceptable prompts

There are two acceptable prompts for LNF: a prompt for when students hesitate and for when they produce letter sounds.

Hesitation Prompt. If the student hesitates for 3 seconds on a letter, score the letter as incorrect, provide the correct letter, point to the next letter, and say:

Keep going.

This prompt may be repeated. For example, if the letters are "p T n" and the student says, "p" then does not say anything for 3 seconds, prompt by saying "**T**", then point to "n" and say:

Keep going.

Repeat this as many times as needed throughout administration. The maximum time for each letter is 3 seconds.

Letter Sound Prompt. If the student provides the letter sound rather than the letter name, say:

Remember, tell me the letter's name, not its sound.

This prompt may be provided *once* during the administration. If the student continues providing letter sounds, mark each letter as incorrect.

Discontinue rules

Discontinue LNF Rule. If the student reads 0 correct letter names within the first line, discontinue LNF, put a bracket after the last letter attempted and record a score of 0.

Discontinue Benchmark Assessments Rule. Benchmark assessment always continues regardless of LNF score.

Scoring rules

LNF provides one score: the number of letters named correctly. Mark student responses according to the rules in the first table below. The second table provides several examples of common situations and how to score in them.

©2026 University of Oregon. All rights reserved.

Correct responses	Do not mark correct responses on the scoring book.
Incorrect responses	Make a slash (/) through each letter named incorrectly.
Self-corrections	If a student makes an error but corrects it within 3 seconds, write "SC" above the letter and score it as correct.

Situation	How to score
Letter reversals	A letter is incorrect if the student substitutes a different letter for the stimulus letter, even if the substituted letter is similar in appearance. (Note that lowercase L does not appear on LNF forms, and the font used in LNF distinguishes the uppercase I from the lowercase L and number 1 very well.)

Letters	Student Says	Scoring Procedure	Correct Letters
b T n E	"d…T…n…E"	~~b~~ T n E	3/4
p S n L	"q…S…m…L"	~~p~~ S ~~n~~ L	2/4
M I k L	"M…L…k…L"	M ~~I~~ k L	3/4

©2026 University of Oregon. All rights reserved.

<table>
<tr><th>Situation</th><th>How to score</th></tr>
<tr><td>Letter sounds</td><td>A letter is incorrect if the student provides the letter-sound for the stimulus letter (e.g., /d/ for “D”). A prompt for providing letter-sounds is allowable only once (see Acceptable Prompts on page 50).
<table>
<tr><th>Letters</th><th>Student Says</th><th>Scoring Procedure</th><th>Correct Letters</th></tr>
<tr><td>b T n E</td><td>“/b/…T…n…E”</td><td>~~b~~ T n E</td><td>3/4</td></tr>
<tr><td>p S n L</td><td>“p…/s/…n…L”</td><td>p ~~S~~ n L</td><td>3/4</td></tr>
<tr><td>M I k L</td><td>“M…I…/k/…L”</td><td>M I ~~k~~ L</td><td>3/4</td></tr>
</table></td></tr>
<tr><td>Omissions</td><td>A letter is incorrect if the student skips the letter. If the student skips an entire line, cross out the line and record a score of 0 for that line.</td></tr>
</table>

LNF Fidelity of Administration

The observer should judge the full test administration. That includes observing setup and directions, timing and scoring the test in parallel with the examiner, checking the examiner’s accuracy in procedures using the fidelity checklist in Appendix D on page 123, and deciding if the examiner passes or needs more practice for each procedure listed.

©2026 University of Oregon. All rights reserved.

Phonemic Segmentation Fluency (PSF)

Applicable grades: Beginning of kindergarten through end of first grade.

Objective: Student breaks words into phonemes for 60 seconds.

Uses: Benchmark and risk assessment; progress monitoring.

Materials

- Scoring book
- There is **NO** student form
- Pen or pencil
- Clipboard
- Timer

Administration

1. Position the clipboard and timer so that the student cannot see what you record.
2. Say these specific directions:

   I am going to say a word. After I say it, you tell me all the sounds you hear in the word. So, if I say the word 'mop', you would say /m/ /o/ /p/.

   If I say the word 'cat' you would say /c/ /a/ /t/.

   Let's try one.

   (1 second pause)

   Tell me the sounds in the word 'sun'.

   Tell me any sounds you hear.

©2026 University of Oregon. All rights reserved.

Student response	Examiner response
CORRECT	
If student says "/s/ /u/ /n/"	**Very good.** **The sounds in "sun" are /s/ /u/ /n/.**
INCORRECT	
Any other response	**The sounds in the word "sun" are /s/ /u/ /n/.** **Your turn.** **Tell me the sounds in "sun." Tell me any sounds you hear.**

OK. Here is your first word.

3. Give the student the first word and start the timer.
4. Follow along in the Scoring Booklet. As the student says the sounds, underline each different, correct, sound segment produced. Put a slash (/) through sounds produced incorrectly. See Acceptable Prompts and Scoring Rules for more details.
5. As soon as the student is finished saying the sounds in the current word, present the next word promptly and clearly.
6. At the end of 60 seconds, stop presenting words and stop the timer. Place a bracket (]) after the last sound provided by the student.

Acceptable prompts

There is only one acceptable prompt for PSF: a prompt for when students hesitate.

Hesitation Prompt. If the student hesitates for 3 seconds, give the next word, and score the word (or remaining sounds in the word if word has been partially segmented) as incorrect by leaving it unmarked (no slashes or underlines). Repeat this prompt as many times as needed throughout administration.

©2026 University of Oregon. All rights reserved.

Discontinue rules

Discontinue PSF Rule. If a student has not given any sound segments correctly in the first 5 words, discontinue PSF, put a bracket after the last word attempted and record a score of 0.

Discontinue Benchmark Assessments Rule. For beginning of kindergarten only, if student does not get any sounds correct in the first 5 words, discontinue PSF and any further benchmark assessments (i.e., NWF and WRF) for that time of year. At all other times of year, benchmark assessment continues regardless of PSF score.

Scoring rules

PSF provides one score: the sum of sound segments produced. Students receive 1 point for each different, correct, part of the word. Mark student responses according to the rules in the first table below. The second table provides several examples of common situations and how to score in them.

Correct responses	Underline the sound segments in the word the student produces that are correctly pronounced.
Incorrect responses	Make a slash (/) through sounds pronounced incorrectly. Circle the item if the student repeats the word correctly, but without segmentation.
Self-corrections	If a student makes an error but corrects it within *3 seconds*, write "SC" above the phoneme and score it as correct.

©2026 University of Oregon. All rights reserved.

Situation	How to score
Schwa sounds	Schwa sounds (/u/) added to consonants are not counted as errors. Some phonemes cannot be pronounced correctly in isolation without a vowel, and some early learning of sounds includes the schwa. For example, if the word is "track," and the student says "tu...ru...a...ku" they would receive 4 of 4 points.

Word	Student Says	Scoring Procedure	Correct Segments
track	"tu...ru...a...ku"	/t/ /r/ /a/ /k/	4/4
bet	"bu...e...tu"	/b/ /e/ /t/	3/3

Situation	How to score
Additions	Additions are not counted as errors if they are separated from the other sounds in the word. For example, if the word is "track," and the student says "t...r...a...ck...s," they would receive 4 of 4 points.

Word	Student Says	Scoring Procedure	Correct Segments
track	"t...r...a...ck...s"	/t/ /r/ /a/ /k/	4/4
top	"s...t...o...p"	/t/ /o/ /p/	3/3
top	"st...o...p"	/~~s~~/ /o/ /p/	2/3
top	"s...t...ol...p"	/t/ /~~o~~/ /p/	2/3
top	"s...t...o...l...p"	/t/ /o/ /p/	3/3

©2026 University of Oregon. All rights reserved.

Situation	How to score
Sound elongation	The student may elongate the individual sounds and run them together as long as it is clear he or she is aware of each sound individually. For example, if the student says, "ssssuuuunnnn," with each phoneme held long enough to make it clear they know the sounds in the word, they would receive credit for 3 phonemes correct. This is a professional judgment and should be based on the student's responses and prior knowledge of the student's instruction. When in doubt, no credit is given.

Word	Student Says	Scoring Procedure	Correct Segments
sun	"ssssuuuunnnn"	/s/ /u/ /n/	3/3

Situation	How to score
Partial segmentation	The student is given credit for each correct sound segment, even if they have not segmented to the phoneme level. Use the underline to indicate the size of the sound segment. For example, if the word is "track," and the student says "tr...ack," they would receive 2 of 4 points.

Word	Student Says	Scoring Procedure	Correct Segments
track	"tr...ack"	/t/ /r/ /a/ /k/	2/4
bet	"b...et"	/b/ /e/ /t/	2/3

©2026 University of Oregon. All rights reserved.

Situation	How to score
Overlapping segmentation	The student receives credit for each different, correct, sound segment of the word. Thus, if the word is "track," and the student says "tra...ack," the student would receive 2 of 4 points because /tra/ and /ack/ are both different, correct, sound segments of "track."

Word	Student Says	Scoring Procedure	Correct Segments
track	"tra...ack"	/t/ /r/ /a/ /k/	2 /4
bet	"be...e...et"	/b/ /e/ /t/	3 /3

Situation	How to score
Mispronounced segment	The student does not receive credit for sound segments that are mispronounced. For example, if the word is "track," and the student says "t...r...a...gs" they would receive no credit for /gs/ because there is no /g/ or /s/ sound segment in the word "track."

Word	Student Says	Scoring Procedure	Correct Segments
track	"t...r...a...gs"	/t/ /r/ /a/ /~~g~~/	3 /4
bet	"p...i...t"	/~~p~~/ /~~i~~/ /t/	1 /3
bet	"d...e...t""	/~~d~~/ /e/ /t/	2 /3

©2026 University of Oregon. All rights reserved.

Situation	How to score
R-Controlled vowels	As discussed earlier in this chapter, r-controlled vowels are technically one phoneme. Students who correctly segment that phoneme or who further segment an r-controlled phonemes into phones should receive full credit. For example, if the word is "car", and the student says "c...ah...r" or "c...ar", they would receive full credit.

Word	Student Says	Scoring Procedure	Correct Segments
car	"c...ah...r"	/k/ /ar/	2/2
car	"c...ar"	/k/ /ar/	2/2
chair	"ch...air"	/ch/ /air/	2/2
chair	"ch...ay...ee...r	/ch/ /air/	2/2

Situation	How to score
No segmentation	If the student repeats the entire word, no credit is given for any sounds. For example, if the word is "track," and the student says "track," circle the entire word and record zero points.

Word	Student Says	Scoring Procedure	Correct Segments
track	"track"	/t/ /r/ /a/ /k/	0/4

Situation	How to score
Spelling	If the student spells the word, no credit is given. For example, if the word is "track," and the student says "t ...r...a...c...k", cross out each sound.

Word	Student Says	Scoring Procedure	Correct Segments
track	"t...r...a...c...k"	/t/ /r/ /a/ /k/	0/4

©2026 University of Oregon. All rights reserved.

Situation	How to score
Omissions	A sound is incorrect if the student omits the sound, but the sound is left unmarked.

Word	Student Says	Scoring Procedure	Correct Segments
track	"tr…" (3 seconds)	/t/ /r/ /a/ /k/	1 /4
bet	"b… t"	/b/ /e/ /t/	2 /3

PSF Fidelity of Administration

The observer should judge the full test administration. That includes observing setup and directions, timing and scoring the test in parallel with the examiner, checking the examiner's accuracy in procedures using the fidelity checklist in Appendix D on page 123, and deciding if the examiner passes or needs more practice for each procedure listed.

©2026 University of Oregon. All rights reserved.

Nonsense Word Fluency (NWF)

Applicable grades: Beginning of kindergarten through end of third grade.

Objective: Student reads or sounds out nonsense words for 60 seconds.

Uses: Benchmark and risk assessment; progress monitoring.

Materials

- Scoring book
- Student form
- Pen or pencil
- Clipboard
- Timer

Administration

1. Position the clipboard and timer so that the student cannot see what you record.
2. Place the student copy of the NWF practice items in front of the student.
3. Say these specific directions:

   Look at this word.

   (point to first word on the practice form)

   It's a make-believe word.

   Watch me read the word: /h/ /a/ /p/, "hap."

   (point to each letter, then run your finger fast beneath the whole word)

   I can say the sounds of the letters, /h/ /a/ /p/

   (point to each letter)

   <u>OR</u> I can read the whole word "hap."

   (run your finger fast beneath the whole word)

   Your turn to read a make-believe word.

   Read this word the best you can.

   (point to the word "lum")

   Make sure you say any sounds you know.

©2026 University of Oregon. All rights reserved.

Student response	Examiner response
CORRECT	
If student says “lum” or “/l/ /u/ /m/”	**That’s right.** **The sounds are “/l/ /u/ /m/” or “lum”.**
INCORRECT	
Any other response	**Remember, you can say the sounds or you can say the whole word.** **Watch me: the sounds are “/l/ /u/ /m/.”** (point to each letter) **Or “lum.”** (run your finger fast beneath the whole word) **Let’s try again.** **Read this word the best you can.** (point to the word “lum”)

(place the student copy of the form in front of the student)

Here are some more make-believe words.

(point to the student form)

Start here

(point to the first nonsense word)

and go across the page

(point across the page)

When I say “Begin,” read the words the best you can.

Point to each letter and tell me the sound <u>OR</u> read the whole word.

Put your finger on the first word.

Ready?

©2026 University of Oregon. All rights reserved.

Begin.

4. Start the timer after saying "**Begin**."
5. Follow along in the Scoring Booklet. As the student says sounds/words, underline each correct sound/word produced. Put a slash (/) through sounds/words produced incorrectly. See Acceptable Prompts and Scoring Rules for more details.
6. At the end of 60 seconds, place a bracket (]) after the last nonsense word for which the student provided sound/word and say, "**Stop**."

Acceptable prompts

There is only one acceptable prompt for NWF: a prompt for when students hesitate. Execution of the prompt depends on whether a student is initially blending nonsense words or sounding them out. If the student is reading words, the rule applies to words; if the student is sounding words out, the rule applies to sounds.

Hesitation Prompt. If student hesitates for 3 seconds on a sound/word, mark the sound/word as incorrect, point to the next sound/word, and say

Keep going.

Repeat this as many times as needed throughout administration. The maximum time for each sound/word is 3 seconds.

Discontinue rules

Discontinue NWF Rule. If a student does not get any sounds correct in the first 5 words, discontinue NWF, put a bracket after the last nonsense word attempted and record a score of 0 for both CLS and WRC.

Discontinue Benchmark Assessments Rule. For middle of kindergarten only, if student does not get any sounds correct in the first 5 words, discontinue NWF and any further benchmark assessments for that time of year (i.e., WRF). At all other times of year, benchmark assessment continues regardless of NWF score.

Scoring rules

NWF provides two scores: the sum of correct letter sounds (CLS) and the sum of words read or recoded correctly (WRC). Every correct letter sound receives 1 point for CLS, regardless of whether

©2026 University of Oregon. All rights reserved.

a student blends. Words read correctly, whether sounded out initially or not, receive 1 point each for WRC. Mark student responses according to the rules in the first table below. The second table provides several examples of common situations and how to score in them.

Correct responses	Underline the letters that the student produces correctly. Underline multiple letters for partially blended words and whole words for fully blended words (with or without sounding out initially).
Incorrect responses	Make a slash (/) through sounds/words produced incorrectly.
Self-corrections	If a student makes an error but corrects it within *3 seconds*, write "SC" above the phoneme and score it as correct.

Situation	How to score
Sounds followed by correct word	When a student sounds out a nonsense word and then blends it, underline the individual letters and then the nonsense word as a whole and score a 3 for CLS and a 1 for WRC. If a student sounds out a nonsense word incorrectly and then blends it correctly, scoring is based on the last utterance and the student would get full credit for both CLS and WRC.

Word	Student Says	Scoring Procedure	Score	
			CLS	WRC
rab	"/r/…/a/…rab"	/r/ /a/ /b/	3/3	1/1
mot	"/m/…/o/…/t/…mot"	/m/ /o/ /t/	3/3	1/1
nar	"/n/…/a/…/p/…nar"	/n/ /a~~r~~/SC	2/2	1/1

©2026 University of Oregon. All rights reserved.

Situation	How to score

Sounds followed by incorrect word

When a student sounds out a nonsense word and then blends it, underline the individual letters and then the nonsense word as a whole and score a 3 for CLS and a 1 for WRC.

If a student sounds out a nonsense word correctly and then blends it **incorrectly,** scoring would be based on the last utterance, and the student would not receive full credit for CLS.

Word	Student Says	Scoring Procedure	Score	
			CLS	WRC
nar	"/n/…/ar/…nap"	/n/ /~~ar~~/	1 /2	0 /1
nar	"/n/…/ar/…/t/…map"	/~~n~~/ /~~ar~~/	0 /2	0 /1
lum	"/l/…/u/…/m/…lim"	/l/ /~~u~~/ /m/	2/3	0/ 1

Repeated sounds

Letter sounds given twice receive credit once. For example, if stimulus word is "rab" and the student says /r/ /a/ /ab/, the student receives only 1 point for the letter sound "a" even though the correct sound was provided twice, and a total CLS score of 3 and a total WRC score of 0.

Word	Student Says	Scoring Procedure	Score	
			CLS	WRC
rab	"r…a…ab"	/r/ /a/ /b/	3 /3	0 /1
mot	"m…o…t…mo…t"	/m/ /o/ /t/	3 /3	0 /1

©2026 University of Oregon. All rights reserved.

Situation	How to score
Partially correct responses	If a word is partially correct, underline the corresponding letters for the sounds produced correctly and word parts for any sounds blended. Put a slash (/) through incorrectly produced letter sounds (to distinguish from omissions; see Omissions scoring rule on page 69). For example, if the word is "rab" and the student says "rayb" (with a long /A/), the letters "r" and "b" would be underlined, and the letter "a" would be slashed with a score of 2 for CLS and 0 for WRC.

Word	Student Says	Scoring Procedure	Score	
			CLS	WRC
rab	"r...ay...b"	/r/ /a/ /b/	2/3	0/1
rab	"rayb"	/r/ /a/ /b/	2/3	0/1
nar	"n...er"	/n/ /ar/	1/2	0/1
nar	"ner"	/n/ /ar/	1/2	0/1

©2026 University of Oregon. All rights reserved.

Situation	How to score
Sounds out of order	Letter sounds produced out of order are scored as incorrect. For example, if the stimulus word is "mot" and the student says /t/ /o/ /m/, only /o/, the letter sound read correctly, would be underlined with a score of 1 for CLS and 0 for WRC. This is true even if the student uses partial or full blending. Blended letter sounds must be correct and in the correct position (beginning, middle, end) to receive credit. If a student reads a nonsense word using blending, letter sounds produced out of order are scored as incorrect.

Word	Student Says	Scoring Procedure	Score	
			CLS	WRC
mot	"t…o…m"	/~~t~~/ /o/ /~~m~~/	1/3	0/1
mot	"to…om…tom"	/~~t~~/ /o/ /~~m~~/	1/3	0/1
mot	"tom"	/~~t~~/ /o/ /~~m~~/	1/3	0/1
mot	"mob"	/m/ /o/ /~~b~~/	2/3	0/1
ag	"ga"	/~~a~~/ /~~g~~/	0/2	0/1

©2026 University of Oregon. All rights reserved.

Situation	How to score
Inserted Sounds	Inserted sounds are not counted against the CLS score but would result in a score of 0 for WRC. This is true whether the insertion is in the beginning, middle or end of a word. For example, if the word is 'com' and the student said 'scom' they would score 3 for CLS and 0 for WRC. If the student is reading whole words, underline the word and include a vertical line where the sound is inserted. An exception to this rule applies when a sound could legitimately be added based on other reasonable pronunciations. In this case a student could receive credit for WRC, even with a sound inserted. For example, the /oo/ sound in words with a long 'U' can be pronounced with or without an additional /y/ sound, as in the difference between 'dune' and 'cute'. Either pronunciation is acceptable and students are not penalized for adding a /y/. Sometimes these exceptions are a result of dialect. For example, with the word-ending 'olk' the 'l' is pronounced in some parts of the country and is silent in other parts of the country. If a student inserts the /l/ sound they would receive full credit for both CLS and WRC.

Word	Student Says	Scoring Procedure	Score	
			CLS	WRC
com	"scom"	/c/ | /o/ /m/	3/3	0/1
com	"crom"	/c/ /o/ /m/	3/3	0/1
hume	"hoom"	/h/ /oo/ /m/	3/3	1/1
hume	"hyoom"	/h/ /oo/ /m/	3/3	1/1
rolk	"roke"	/r/ /O/ /k/	3/3	1/1
rolk	"rolk"	/r/ /O/ /k/	3/3	1/1

©2026 University of Oregon. All rights reserved.

<table>
<tr><th>Situation</th><th>How to score</th></tr>
<tr><td>R-Controlled Vowels</td><td>Vowels followed by an ‘r’ are counted as one phoneme. However, if a student separates the vowel sound from the /r/ sound, they are not penalized, as long as this does not substantially distort the sound made by the r-controlled vowel. For example, the word ‘nar’ has two sounds: /n/ and /ar/. If a student said “n...ah...r” they would still score 2 for CLS and 0 for WRC.
<table>
<tr><th rowspan="2">Word</th><th rowspan="2">Student Says</th><th rowspan="2">Scoring Procedure</th><th colspan="2">Score</th></tr>
<tr><th>CLS</th><th>WRC</th></tr>
<tr><td>nar</td><td>“n...ar”</td><td>/n/ /ar/</td><td>2/2</td><td>0/1</td></tr>
<tr><td>nar</td><td>“n...ah...r”</td><td>/n/ /ar/</td><td>2/2</td><td>0/1</td></tr>
<tr><td>nair</td><td>"n...air"</td><td>/n/ /air/</td><td>2/2</td><td>0/1</td></tr>
<tr><td>nair</td><td>"n...aye...r"</td><td>/n/ /air/</td><td>2/2</td><td>0/1</td></tr>
</table></td></tr>
<tr><td>Omissions</td><td>If a student skips a word or row, this can be marked by leaving the sound/word blank or using a slash and move to the next word, row, or page with the student.</td></tr>
</table>

NWF Fidelity of Administration

The observer should judge the full test administration. That includes observing setup and directions, timing and scoring the test in parallel with the examiner, checking the examiner’s accuracy in procedures using the fidelity checklist in Appendix D on page 123, and deciding if the examiner passes or needs more practice for each procedure listed.

©2026 University of Oregon. All rights reserved.

Word Reading Fluency (WRF)

Applicable grades: Beginning of kindergarten through end of third grade.

Objective: Student reads words for 60 seconds.

Uses: Benchmark and risk assessment; progress monitoring.

Materials

- Scoring book
- Student form
- Pen or pencil
- Clipboard
- Timer

Administration

1. Position the clipboard and timer so that the student cannot see what you record.
2. Place the student copy of the WRF form in front of the student.
3. Say these specific directions:

   Please read from this list of words.

   (point to the student form)

   Start here

   (point to the first word)

   and go across the page.

   (point across the page)

   When I say "Begin," point to each word and read it the best you can.

   If you get stuck, I will tell you the word, so you can keep reading.

   Put your finger on the first word.

   Ready?

   Begin.

4. Start the timer when student says first word.
5. Follow along in the Scoring Booklet. As the student provides responses, put a slash (/) through each word read incorrectly. See Acceptable Prompts and Scoring Rules for more

©2026 University of Oregon. All rights reserved.

details.

6. At the end of 60 seconds, place a bracket (]) after the last word read and say, "**Stop**."

Acceptable prompts

There is only one acceptable prompt for WRF: a prompt for when students hesitate.

Hesitation Prompt. If student hesitates for 3 seconds on a word, give the correct word, mark the word as incorrect, point to the next word, and say:

Keep going.

Repeat this as many times as needed throughout administration. The maximum time for each word is 3 seconds.

Discontinue rules

Discontinue WRF Rule. If a student does not get any words correct in the first line (5 words), discontinue WRF, put a bracket after the last word attempted and record a score of 0.

Discontinue Benchmark Assessments Rule. For beginning of first grade only, if student does not get any words correct in the first 5 words: discontinue WRF and any further benchmark assessments for that time of year (i.e., ORF). At all other times of year, benchmark assessment continues regardless of WRF score.

Scoring Rules

WRF provides one score: the sum of words read correctly. Mark student responses according to the rules in the first table below. The second table provides several examples of common situations and how to score in them.

Correct responses	Do not mark correct items on the scoring book.
Incorrect responses	Put a slash (/) through words produced incorrectly.
Self-corrections	If a student makes an error and corrects it within *3 seconds*, write "SC" above the word and score it as correct.

©2026 University of Oregon. All rights reserved.

<table>
<tr><th>Situation</th><th>How to score</th></tr>
<tr><td>Sounded out words</td><td>If a word is sounded out without blending, it is incorrect. If a word is sounded out and then blended, it is correct.
<table>
<tr><th>Words</th><th>Student Says</th><th>Scoring Procedure</th><th>Score</th></tr>
<tr><td>joy draw cloud</td><td>“/j/ /oy/ draw cloud”</td><td>joy draw cloud</td><td>2 /3</td></tr>
<tr><td>joy draw cloud</td><td>“/j/ /oy/ joy draw cloud”</td><td>joy draw cloud</td><td>3 /3</td></tr>
</table></td></tr>
<tr><td>Word order</td><td>Words read correctly but in the wrong order are scored as incorrect.
<table>
<tr><th>Words</th><th>Student Says</th><th>Scoring Procedure</th><th>Score</th></tr>
<tr><td>joy draw cloud</td><td>“joy cloud draw”</td><td>joy draw cloud</td><td>1 /3</td></tr>
</table></td></tr>
<tr><td>Omissions</td><td>A word is incorrect if the student skips the word. If the student skips an entire line, cross out the line and record a score of 0 for that line.</td></tr>
</table>

WRF Fidelity of Administration

The observer should judge the full test administration. That includes observing setup and directions, timing and scoring the test in parallel with the examiner, checking the examiner’s accuracy in procedures using the fidelity checklist in Appendix D on page 123, and deciding if the examiner passes or needs more practice for each procedure listed.

©2026 University of Oregon. All rights reserved.

Oral Reading Fluency (ORF)

Applicable grades: Beginning of first grade through end of eighth grade.

Objective: Student reads a passage aloud for 60 seconds.

Uses: Benchmark and risk assessment; progress monitoring.

Materials

- Scoring book
- Student form
- Pen or pencil
- Clipboard
- Timer

Administration

1. Position the clipboard and timer so that the student cannot see what you record.
2. Place the student copy of the ORF form in front of the student.
3. Say these specific directions:

   Please read this

   (point to the 1st word of the 1st paragraph of the passage)

   out loud.

   If you get stuck, I will tell you the word, so you can keep reading.

   When I say "Stop" I may ask you to tell me about what you read, so do your best reading.

   Start here

   (point to the first word of the passage).

   Ready?

   Begin.

4. Start the timer when the student says the first word of the passage. Do NOT count the title. If the student fails to say the first word after 3 seconds, tell the student the word and mark it as incorrect, then start the timer.

©2026 University of Oregon. All rights reserved.

5. Follow along in the Scoring Booklet. As the student provides responses, put a slash (/) through each word read incorrectly. See Acceptable Prompts and Scoring Rules for more details.
6. At the end of 60 seconds, place a bracket (]) after the last word read and say, "**Stop**."

Acceptable prompts

There is only one acceptable prompt for ORF: a prompt for when students hesitate.

Hesitation Prompt. If student hesitates for 3 seconds on a word, give the correct word, and mark the word as incorrect. Repeat this as many times as needed throughout administration. The maximum time for each word is 3 seconds.

Discontinue rules

Discontinue ORF Rule. If the student does not read any words correctly in the first line of the passage, discontinue ORF, put a bracket after the last word attempted and record a score of 0.

Discontinue Benchmark Assessments Rule. Benchmark assessment always continues regardless of ORF score.

Scoring rules

ORF provides two scores: the sum of words read correctly and an accuracy percentage. The accuracy percentage is calculated by dividing the sum of words read correctly by the number of total words attempted (including errors) and multiplying by 100:

$$\textit{Accuracy} = \frac{\textit{words read correctly}}{\textit{total words read}} \times 100$$

Mark student responses according to the rules in the first table below. The second table provides several examples of common situations and how to score in them.

©2026 University of Oregon. All rights reserved.

Correct responses	Do not mark correct items on the scoring book.
Incorrect responses	Put a slash (/) through words produced incorrectly.
Self-corrections	If a student makes an error and corrects it within *3 seconds*, write "SC" above the word and score it as correct.

Situation	How to score
Insertions	Inserted words are ignored and not counted as errors. The student does not get points for inserted words. If the student frequently inserts extra words, it may be worth noting the pattern at the bottom of the scoring page.

Passage	Student Says	Scoring Procedure	Score
I have a dog.	"I have a new dog."	I have a dog.	4/4
The walk was fun.	"The walk was really fun."	The walk was fun.	4/4

Repetitions	Words that are repeated are not scored as incorrect so long as they are read correctly. They are treated as insertions and ignored in scoring.

Passage	Student Says	Scoring Procedure	Score
I have a dog.	"I have a ... I have a dog."	I have a dog.	4/4

©2026 University of Oregon. All rights reserved.

Situation	How to score

Sounded out words

A word is scored as incorrect if it is sounded out correctly but not blended. If it is blended, it is scored as correct.

Passage	Student Says	Scoring Procedure	Score
We like to read.	"We like to rrrr … eeee … d read."	We like to read.	4 /4
We like to read.	"We like to rrrr … eeee … d."	We like to ~~r~~ead.	3 /4

Abbreviations

Abbreviations should be read in the way you would normally pronounce the abbreviation in conversation. For example, ASAP could be read as "ay ess ay pea" or "ay sap" and Dr. would be read as "doctor."

Passage	Student Says	Scoring Procedure	Score
Tell me ASAP.	"Tell me ay ess ay pea."	Tell me ASAP.	3 /3
Tell me ASAP.	"Tell me ay sap."	Tell me ASAP.	3 /3
Dr. Jones looked at my teeth.	"Doctor Jones looked at my teeth."	Dr. Jones looked at my teeth.	6 /6
Dr. Jones looked at my teeth.	"'D' 'r' Jones looked at my teeth.	~~Dr~~. Jones looked at my teeth.	5 /6

©2026 University of Oregon. All rights reserved.

<table>
<tr><th>Situation</th><th>How to score</th></tr>
<tr><td>Mispronounced words</td><td>A word is scored as incorrect if it is pronounced incorrectly in the context of the sentence.
<table>
<tr><th>Passage</th><th>Student Says</th><th>Scoring Procedure</th><th>Score</th></tr>
<tr><td>We like to read.</td><td>“We like to red.”</td><td>We like to read.</td><td>3/4</td></tr>
</table></td></tr>
<tr><td>Word order</td><td>All words that are read correctly but in the wrong order are scored as incorrect.
<table>
<tr><th>Passage</th><th>Student Says</th><th>Scoring Procedure</th><th>Score</th></tr>
<tr><td>The green park has flowers.</td><td>“The park green has flowers.”</td><td>The green park has flowers.</td><td>3/5</td></tr>
</table></td></tr>
<tr><td>Omissions</td><td>Omitted words are scored as incorrect. If a student skips an entire row, cross out the row and mark the skipped words incorrect.</td></tr>
</table>

ORF Fidelity of Administration

The observer should judge the full test administration. That includes observing setup and directions, timing and scoring the test in parallel with the examiner, checking the examiner’s accuracy in procedures using the fidelity checklist in Appendix D on page 123, and deciding if the examiner passes or needs more practice for each procedure listed.

©2026 University of Oregon. All rights reserved.

Maze

Applicable grades: Beginning of second grade through end of eighth grade.

Objective: Student silently reads a passage for 180 seconds, choosing the best multiple-choice answer for missing words.

Uses: Benchmark and risk assessment; progress monitoring.

Materials

- Maze administration directions and scoring key
- Student worksheets (one per student)
- Pen or pencil (one per student)
- Clipboard
- Timer

Administration

1. Say:

   I am going to give you a worksheet. When you get your worksheet, please write your name at the top and put your pencil down.

2. Hand out the Maze student worksheets.
3. Make sure students have written their names down before proceeding.
4. Say these specific directions:

   You are going to read a passage with some words missing from it. For each missing word you will see a box with three words in it. Your job is to circle the word you think makes the most sense in the context of the passage. Let's look at the Practice Passage together. Listen as I read.

   (pause)

   Tom goes to a school far from his house. Every morning, he takes a school

   (pause)

   art, bus, work

©2026 University of Oregon. All rights reserved.

(pause)

to go to school.

(pause)

Let's stop there. Let's circle the word "bus" because I think "bus" makes the most sense here. Listen to how that sentence sounds now.

Every morning, he takes a school bus to go to school.

Now it's your turn. Read the next sentence silently to yourself.

When you come to a box, read all the words in the box and circle the word that makes the most sense to you. When you are done, put your pencil down.

5. Allow up to 30 seconds for students to complete the example and put their pencils down. If necessary, after 30 seconds say **Put your pencil down**.
6. As soon as all students have their pencils down, say

Good job.

Now listen. In the

(pause)

afternoon, library, morning,

(pause)

he also takes a bus home. You should have circled "afternoon" because "afternoon" makes the most sense.

(pause)

Listen. In the afternoon, he also takes a bus home.

Okay, when I say "Begin," turn the page and start reading the passage silently. Start on the page with the title. When you come to a box, read all the words in the box and circle the word that makes the most sense in the passage.

You will stop when you come to a stop sign or I say Stop.

©2026 University of Oregon. All rights reserved.

Ready?

Begin.

7. Start the timer.
8. At the end of 3 minutes, stop the timer and say, **Stop. Put your pencils down**.
9. Make sure all students have stopped working and collect all the student worksheets.

Acceptable Prompts

There are three acceptable prompts for Maze: one for when students read aloud, another for when students skip pages, and the other for when students stop working. These prompts can be used as often as necessary.

Student Reading Aloud Prompt. If a student reads the passage out loud, say:

Please read the passage silently.

Student Skipping Pages Prompt. If a student skips an entire page, say:

Please be sure not to skip pages.

Student Stopped Working Prompt. If a student stops working, say:

Please keep going until I tell you to stop. Just do your best work.

Discontinue rules

There are no discontinue rules for Maze. Every student should be encouraged to try their best until three minutes have passed.

Scoring rules

Maze provides one score that is derived by summing up the number of items answered correctly and subtracting one-half the sum of items answered incorrectly. Worksheets are scored after the assessment has been completed, and students are not present. Use the scoring key to mark answers as correct or incorrect.

1. A response is correct if the student clearly circled or otherwise marked (e.g., underlined or checked) the correct word.
2. Mark a slash (/) through any incorrect responses. Incorrect responses include situations when the wrong answer is circled or otherwise marked, more than one answer is marked, or an item is left blank (only if it occurs before the final item answered).

©2026 University of Oregon. All rights reserved.

3. If there are erasure marks, scratched out words, or any other extraneous markings, and the student's final response is obvious, score the item based on the final response.
4. Items left blank after the last response are not slashed or counted as incorrect.
5. Count the number of items answered that are not slashed to obtain the number of items answered correctly. Enter the total next to the word Correct on the student's booklet.
6. Count the number of items marked with a slash. Enter the total next to the word Incorrect on the student's booklet.
7. Calculate the adjusted score (unnecessary for DIBELS Data System and Amplify customers) using the following formula:

$$Maze = Correct - \frac{Incorrect}{2}$$

By definition, this formula will sometimes result in scores with decimal values. **Decimal values should not be rounded, but negative numbers should be rounded to zero.**

Mark student responses according to the rules in the first table below. The second table provides several examples of common situations and how to score in them.

Correct responses	Do not mark correct items.
Incorrect responses	Put a slash (/) through items answered incorrectly, skipped (before the last valid response), or marked in a confusing manner.
Self-corrections	If a student makes corrections to a response, the answer is counted as correct so long as the final intended answer is both clear and correct.

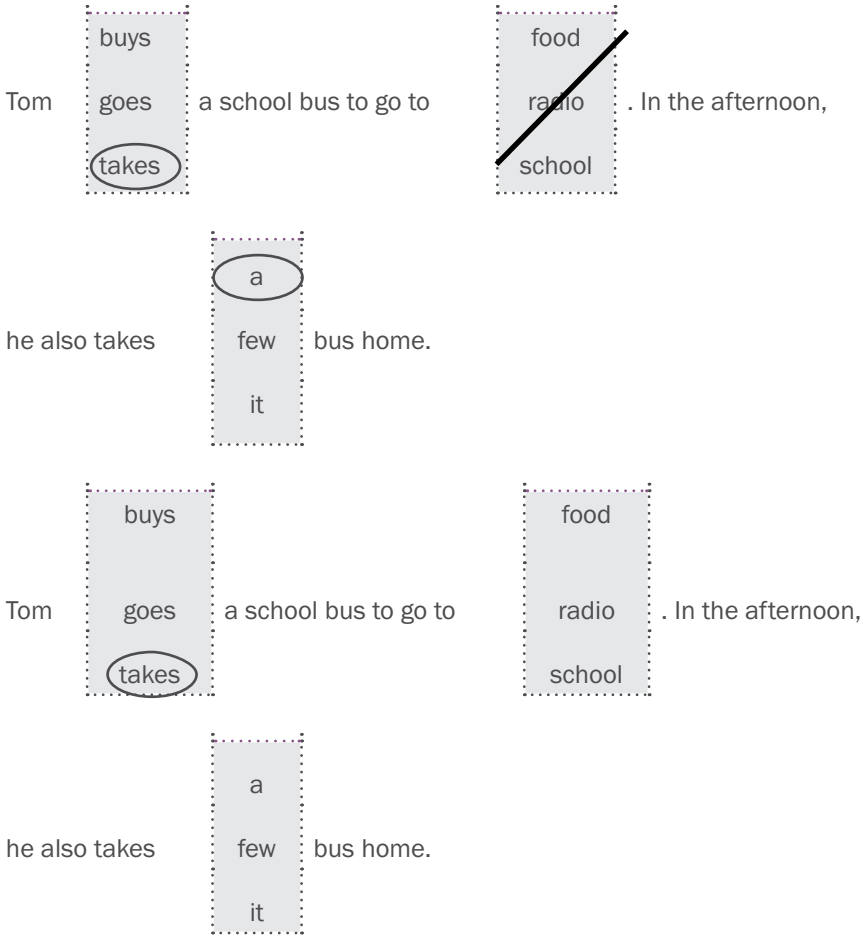
©2026 University of Oregon. All rights reserved.

Situation	How to score
Inconsistent marking	Students sometimes change how they mark the correct answer. So long as the student's intention is clear and correct, changes in marking system are not penalized. In the example below, the student gets 3 items correct and none incorrect.

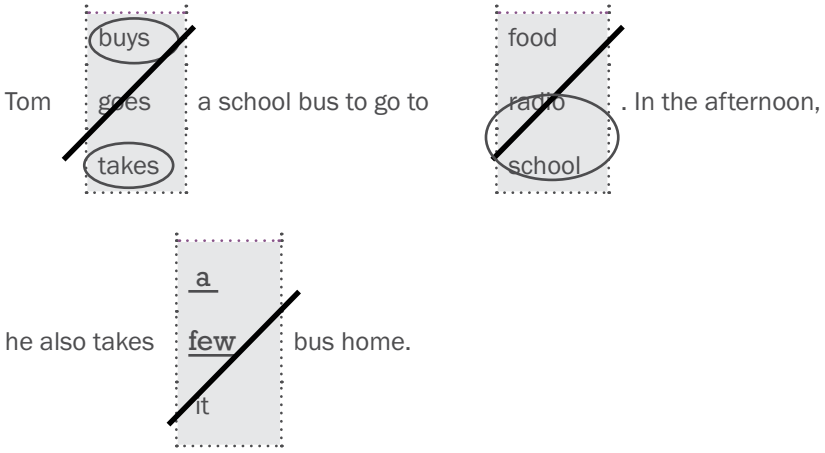
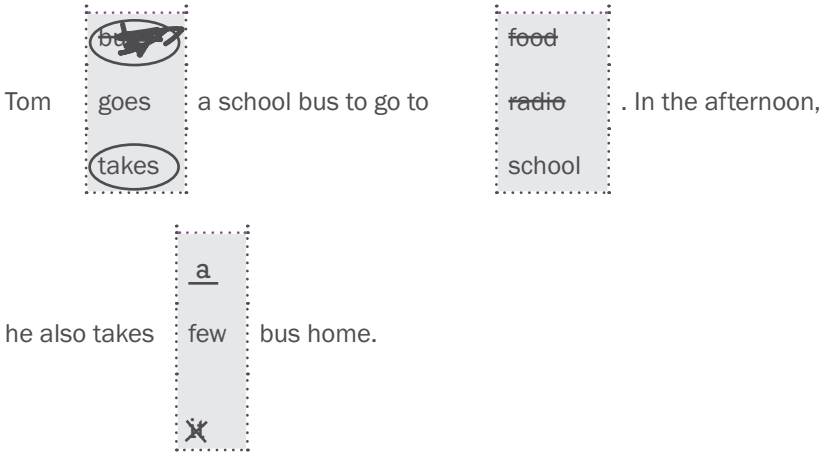


©2026 University of Oregon. All rights reserved.

Situation	How to score
Skipped items	Skipped items are marked incorrect when they are clearly skipped (i.e., a later item is answered), as in the first example below. They are left unmarked and not counted as correct or incorrect if no subsequent item is answered, as in the second example below. In the first example, the student gets 2 correct and 1 incorrect. In the second example, the student gets 1 correct and none incorrect.



©2026 University of Oregon. All rights reserved.

Situation	How to score
Unclear or multiple responses	When more than one choice is marked, and the intended final answer is not clear, the item is scored as incorrect. In the example below, the child gets no items correct and 3 incorrect. 
Multiple marks with clear intention	An item is scored as correct even in the presence of multiple marks if the final intention is clear and correct. In the example below, the child gets 3 items correct and none incorrect. 

©2026 University of Oregon. All rights reserved.

Maze Fidelity of Administration

The observer should judge the full test administration and subsequent scoring and calculations. That includes observing setup and directions, timing and scoring the test in parallel with the examiner, checking the examiner's accuracy in procedures using the fidelity checklist in Appendix D on page 123, and deciding if the examiner passes or needs more practice for each procedure listed.

©2026 University of Oregon. All rights reserved.

Chapter 3: Interpreting DIBELS 8th Edition Scores

This chapter covers the interpretation of DIBELS 8th Edition scores. Topics include the scores available for DIBELS 8 and cautions in interpreting results. Please see the technical report for more information about how various derived scores were developed.

DIBELS 8 offers five types of scores: composite scores, risk classifications, percentile ranks, raw scores, and Zones of Growth. These scores offer teachers a wealth of information that can be used in planning instruction and monitoring student growth. Each is discussed in turn. For information about DIBELS 8 benchmark goals, please refer to: https://dibels.uoregon.edu/materials.

DIBELS 8 Composite Scores

The composite should be the primary score used to determine risk and needed support. It is the most reliable and valid score in the DIBELS 8 suite of assessments. DIBELS 8th Edition provides composite scores as a means of interpreting and reporting student performance across subtests. The approach to creating the composite scores represents a marked improvement over the DIBELS Next approach in that a confirmatory factor analysis (CFA) was used to determine the optimal weighting of DIBELS subtest raw scores while simultaneously accounting for relations among subtests. Our primary concern was correcting for the fact that NWF and ORF each contribute two scores to the composite. These analyses are described in greater detail in the Appendix F on page 136. The final CFA models for kindergarten through third grade utilized all available subtests and accounted for the covariance between NWF scores. The final CFA models for fourth through eighth grade utilized all available subtests without accounting for covariances. All solutions were scaled based on the initial norming sample so that 360 represented the mean at the beginning of the year, 400 represented the mean at the middle of the year, and 440 represented the mean at the end of the year for that sample, with 40 as the standard deviation for each time of year. Note that expected growth for an academic year is determined not by these numbers but by the updated norms that are released every few years. For example, based on the 2024-25 normative sample of approximately 890,000 Grade 1 students from all 50 US states, the average Composite score was 339 at the beginning of the year, and 469 at the end of the year, resulting in average Composite score growth of 130 points. The percentile norms can be found on the

©2026 University of Oregon. All rights reserved.

Research page of the DIBELS website: https://dibels.uoregon.edu/research.

DIBELS 8 Risk Classification

Although raw scores are not ideal for tracking growth, they can be utilized for screening purposes. Specifically, we created cut-scores for determining students' risk using raw scores. To support this use, we have provided three types of cut-scores for classifying students. Cut- scores are summarized by grade, measure, and time of year in Appendix E on page 130.

The first score, called the risk cut-score, can be used to classify students who are well below benchmark in their performance and at risk for reading difficulties, including dyslexia. On average, the at-risk cut-score identifies 80% of students performing below the 20th percentile on an external outcome measure at the end of the year. Students falling below this cut-score are designated with the color red in the DIBELS 8 benchmark documentation.

The second score, called the benchmark goal, can be used to classify students who are performing at benchmark levels and are at minimal risk and on track for meeting grade-level proficiency goals from those who are below benchmark performance levels and thus at some risk for not meeting proficiency goals.

On average, this cut-score identifies 80% or more of students performing below the 40th percentile rank on an external measure of reading ability at the end of the year. Students falling above this cut-score are typically in need of core support alone, meaning the general curriculum should serve these children well. Students falling between the risk and benchmark cut-scores are at some risk for not meeting proficiency goals compared to those who are on track for meeting proficiency goals. These students are in need of strategic support. Students falling below this cut- score but above the risk cut-score are designated with the color yellow in the DIBELS 8 benchmark documentation.

Finally, we have introduced a third cut-score, which represents an ambitious goal for students, and can be used to classify students who are performing well above benchmark and are at negligible risk for not meeting proficiency goals. The ambitious cut-score is designed to identify the students who are least at risk in reading. Although students scoring above the benchmark goal are generally at minimal risk, the ambitious goal cut-score provides a second means of determining how secure a student's likelihood of success is. Because the ambitious cut was determined by maximizing sensitivity

©2026 University of Oregon. All rights reserved.

(which is a statistic expressing the percentage of students falling below a specified score that a given cut-score identifies), students who score at or above this cut are at truly negligible risk of scoring below the 40th percentile rank on a criterion reading measure at the end of the school year.

In this case, negligible can be interpreted as meaning 0-10% of students who scored at or above the ambitious cut scored below the 40th percentile rank. In other words, students scoring above the ambitious cut-score have a strong likelihood of performing at an average or above average level for their grade at the end of the year. Students falling below this cut-score but above the benchmark cut- score are designated with the color green in the DIBELS 8 benchmark documentation, while those falling above this cut are designated with the color blue. Students who fall at or above the ambitious cut-score have a greater chance of performing above the 40th percentile rank on an external measure of reading ability at the end of the year than do students who fall between the benchmark and ambitious cut-scores. Students falling above this cut-score are very likely in need of core support alone, meaning the general curriculum should serve these children well. Students performing well above benchmark may benefit from instruction on more advanced skills.

DIBELS 8 Percentile Ranks

Percentile ranks (also known as percentiles) are a way of expressing student performance relative to the norming sample for DIBELS 8. Percentiles look like percentages and represent the percentage of the norming sample that a given student scores at or above on a given subtest. For example, a student who is at the 60th percentile scored the same as or higher than 60% of the norming sample. Because DIBELS researchers made strong efforts to recruit a nationally representative sample when norming DIBELS 8, percentile ranks have strong generalizability.

DIBELS 8 Raw scores

Raw scores are the most basic score available. They generally represent the number of items a student has answered correctly in one minute, with a few exceptions. Maze provides an adjusted raw score where half the number of incorrect items is subtracted from the total number correct. ORF Accuracy is the proportion of words read correctly in one minute and is derived by dividing the number of words read correctly by the total number of words read, including those that were incorrect.

©2026 University of Oregon. All rights reserved.

Raw scores have weaknesses in their interpretation. Despite strenuous efforts to create equivalent forms, differences in difficulty between forms still occur. While these "form effects" are generally quite mild for many subtests as a result of the constrained item pool (e.g., LNF), they become more apparent in subtests involving connected text (i.e., ORF and Maze). Where form effects are more pronounced, differences in scores over time can be obscured or exaggerated. For example, a student who scores 100 words-correct-per-minute (WCPM) in the beginning of year and 120 WCPM in the middle of year has indeed read the middle of year passage at a faster rate, but whether the difference in 20 WCPM is due to actual growth or the middle of year passage simply being easier to read remains unclear. Because form effects can make interpreting student progress difficult, DIBELS 8 offers several alternative score types as described in the foregoing sections.

DIBELS 8 Zones of Growth®

DIBELS 8th Edition also offers scores that can be used to interpret growth relative to the norming sample by defining percentile gains, which are normative data regarding changes in performance over time. Percentile gains facilitate comparisons of an individual student's performance over time relative to the performance over time of other students with a similar starting score. These comparisons provide a more nuanced understanding of student progress than cut-scores or percentile ranks. They are an especially useful tool for evaluating the progress of students who perform below the benchmark level and whose performance over time needs to be monitored more closely. Zones of Growth are only available when using a data system like the DIBELS Data System or mCLASS through Amplify.

DIBELS 8th Edition and Dyslexia Screening

DIBELS 8th Edition features revised versions of LNF, PSF, and NWF that improve their ability to screen for reading-related weaknesses commonly associated with dyslexia, such as rapid naming, phonological awareness, and the alphabetic principle. Specifically, LNF was adapted to improve its validity as a rapid naming measure, PSF was adapted to improve its evidence as a more general phonological awareness measure, and NWF was adapted to better represent the alphabetic principle. DIBELS measures have increasingly been identified by states as measures that can meet new legislated dyslexia screening mandates across the country. Thus, our revisions seek to provide states and schools with the evidence they need to feel confident that DIBELS can fill that purpose. The validity chapter of

©2026 University of Oregon. All rights reserved.

our technical report highlights where evidence supports dyslexia screening.

Nevertheless, DIBELS 8 measures are not a substitute for a complete diagnostic assessment. DIBELS 8 is designed to offer educators an efficient way to screen all of their students for risk in critical areas and more importantly to direct support where it is most needed. For those seeking to use DIBELS 8 to comply with dyslexia screening requirements, we recommend following your state's guidelines for screening. For those without state guidelines, risk on LNF and PSF subtests in kindergarten and first grade and NWF in first through third grade could be used to understand potential risk for dyslexia.

It is important to recognize that these tools are intended to screen for risk and do not render diagnosis regarding dyslexia. While DIBELS measures effectively capture most students with true reading difficulty, many students who are flagged may not prove to meet criteria for dyslexia diagnosis on a more comprehensive evaluation protocol.

As a result, we encourage educators to use DIBELS information primarily to guide their early intervention services and to match students with the appropriate type and level of instruction. All students, including those with dyslexia, can benefit tremendously from effective instruction in phonological awareness and the alphabetic principle, particularly when it is provided early in their academic development. Such support is facilitated through DIBELS 8's identification of students at risk for difficulty in key skill areas during the earliest, critical opportunities for intervention.

Cautions in Interpreting DIBELS 8 Scores

Even though DIBELS 8th Edition has undergone rigorous research and development procedures, no test is ever 100% reliable and accurate. Moreover, no single test should drive high-stakes decisions made about individual students. DIBELS 8 is not a diagnostic measure in the sense that it cannot diagnose the root causes of reading problems, although using all the subtests provided within a grade can lead to strong hypotheses. Nonetheless, hypotheses regarding the origins and diagnosis of reading problems should be interpreted with caution and tested through the use of other measures and observations. Beyond this general caution, which applies to any single test, there is one additional caution worthy of mention, namely inter-rater reliability.

The reliability statistics reported in the DIBELS 8 Technical Manual were obtained after teachers were well trained in the administration and scoring of DIBELS 8. Although we obtained excellent inter-

©2026 University of Oregon. All rights reserved.

rater reliability during the course of DIBELS 8 research, we do not report it in this manual. Inter-rater reliability obtained in a study has no bearing on the use of a measure in schools other than the fact that it suggests high inter-rater reliability is possible to achieve. In other words, the reliability of different raters cannot be assumed and should be established in the specific context in which DIBELS 8 is to be used. In addition to initially training test administrators and assessing inter-rater agreement, DIBELS 8 users should recalibrate (i.e., assess inter-rater agreement after a certain period and retrain as needed) at least once a year.

©2026 University of Oregon. All rights reserved.

Chapter 4: Progress Monitoring with DIBELS 8th Edition

In this section, we discuss the specific use of DIBELS 8th Edition for monitoring student progress. Topics include recommendations regarding which subtest to use, frequency of progress monitoring, and decisions regarding when to monitor off-grade level and when to change progress monitoring subtest or intervention.

Choosing a Subtest for Progress Monitoring

One critical step in progress monitoring students who receive intervention is knowing which subtest to use. Generally speaking, best practice involves monitoring progress for the skill on which intervention is most focused. In no situation should student progress be monitored with a subtest on which they did not demonstrate risk, and LNF should never be used for progress monitoring. Nonetheless, many students will have multiple indicators of risk and receive multi-component interventions. Note that it may be advisable for students receiving multi-component interventions to have their progress monitored on more than one subtest. However, we offer guidelines for how to pick a single progress monitoring subtest to use under specific conditions.

As with prior editions of DIBELS, NWF and ORF are the strongest measures for capturing change over time. As a result, we recommend preferencing these subtests for progress monitoring in the grades in which these subtests are available and where a student has demonstrated risk on one of these subtests.

Beyond this general recommendation, we also suggest that the subtest used for progress monitoring be aligned to the focus of intervention for a student. Therefore, a student who is receiving intervention focused solely on phonological awareness, but not decoding, would be best progress-monitored with PSF. Similarly, we would recommend using WRF for a student who is receiving intervention focused on improving sight word recognition, but who is a strong decoder. However, we maintain that NWF and ORF are the best ways to monitor progress for most children.

As with previous editions, progress monitoring with LNF is not supported. LNF is best used as a risk indicator because it is not a foundational skill in reading in the same way that other DIBELS subtests are.

©2026 University of Oregon. All rights reserved.

For Maze, infrequent progress monitoring is recommended because reading comprehension does not improve rapidly enough, even with intensive intervention, to be observable after short intervals (e.g., Deno et al., 2008; Espin, Wallace, Lembke, Campbell, & Long, 2010; Shin, Deno, & Espin, 2000). Thus, we recommend Maze for progress monitoring no more than two to three times between benchmark occasions. Nonetheless, we offer enough forms to support progress monitoring up to five times between benchmark occasions.

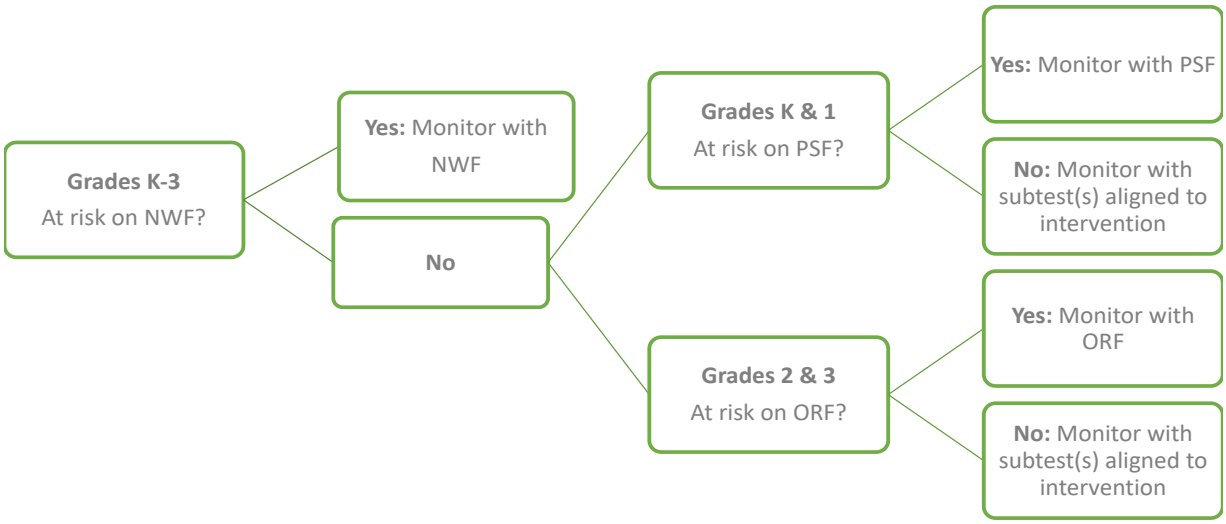


Figure 4.1 Decision tree for choosing a progress monitoring subtest.

Frequency of Progress Monitoring

Another important step in progress monitoring students who receive intervention is knowing how often to monitor progress. In general, the more foundational the skill and the more intensive the intervention, the more frequent progress monitoring should be. However, it is rarely advisable to progress monitor more than once a week. In fact, to avoid excessive assessment, we recommend progress monitoring every two weeks in kindergarten through third grade. Depending on the intensity of intervention, progress monitoring could be as frequent as every second or third week in Grades 4 and up. The additional elapsed time in these grades is recommended based on the slower ORF growth typically observed in the upper grades (e.g., Christ, Silberglitt, Yeo, & Cromier, 2010; Nese,

©2026 University of Oregon. All rights reserved.

Biancarosa, Anderson, Lai, Alonzo, & Tindal, 2012; Nese, Biancarosa, Cummings, Kennedy, Alonzo, & Tindal, 2013). For students at some risk and receiving strategic support, progress monitoring every four weeks (or four to five weeks in Grade 4 and up) is generally adequate.

The exception to these guidelines is progress monitoring with Maze. Unfortunately, even under intensive intervention, reading comprehension is difficult to improve rapidly. As a result, we recommend that progress monitoring with Maze occur no more than once to twice between benchmark periods (i.e., monthly assessment; e.g., Deno et al., 2008; Espin et al., 2010; Shin et al., 2000). Nonetheless, beginning in 2021, we provide 10 progress monitoring forms per grade, allowing for more frequent progress monitoring where necessary.

Table 4.1. Recommended Progress Monitoring Frequency

Grades	Subtests	At Risk (red)	At Some Risk (yellow)
K-3	PSF, NWF, WRF, ORF	Every 2 weeks	Every 4 weeks
4-8	ORF	Every 2-3 weeks	Every 4-5 weeks
2-8	Maze	Up to 2 times between benchmarks	Up to 2 times between benchmarks

Determining Response to Intervention

A challenging aspect of progress monitoring students who receive intervention is knowing how to judge whether a student is responding to intervention. In the past, researchers have recommended as many as eight or more assessment occasions before deciding (e.g., Christ, 2006; Christ, Zopluoglu, Long, & Monaghen, 2012). However, sixteen weeks, if monitoring is conducted as recommended, is simply too long to wait to determine if very vulnerable learners are responding to intervention. Moreover, the research literature relies on model-based estimates of growth, which is not how schools analyze data to make decisions about students at risk.

©2026 University of Oregon. All rights reserved.

As a result, we base DIBELS 8 recommendations for determining response to intervention relative to goals for student growth. Specifically, we recommend setting an end of year goal for a student, where the default goal will typically be the end of year benchmark cut-score. When graphed, student scores on progress monitoring administrations should be mapped relative to an aimline, which is drawn from the benchmark result that occasioned intervention to the end of year goal. So long as student scores hew closely to or above the aimline, the student shows signs of response to intervention.

However, if a student obtains four consecutive data points below the aimline, it indicates the need for either a change in intervention or, in Grades 1 and up, the potential need for off-grade level monitoring (see next section). This guidance applies to all progress monitoring subtests except ORF. Due to the exceptional reliability of ORF and its more frequent use in upper grade levels, only three data points below the aimline are required for this determination when ORF is the progress monitoring subtest.

We do not offer guidance for discontinuing intervention altogether. That decision will depend on a combination of how far a student has progressed, local need, and resources available. However, if a student reaches the end of year benchmark goal on a subtest, it is common-sense to discontinue intervention at that point.

Off-grade Progress Monitoring

One additional challenge in progress monitoring students who receive intervention is determining when students are so far below grade level that progress monitoring is best conducted using off-grade-level forms. For Grade 1 to 8 students who begin the year at or below the 10th percentile rank based on national DIBELS 8 norms, schools may want to consider progress monitoring with an off-grade-level form, especially for older students who have a history of risk. Begin by going one grade down and go further down if needed using the same criteria (i.e., at or below the 10th percentile rank for the new grade). More conservatively, the decision to move off-grade level for progress monitoring will rely on the guidelines expressed in the previous section. That is, when three or more ORF observations or four or more observations on other subtests fall below the aimline, either a change in intervention or off-grade-level monitoring is necessary. You can find the 10th percentile document on the Research page on the DIBELS website: https://dibels.uoregon.edu/research.

©2026 University of Oregon. All rights reserved.

For benchmarking students, always use on-grade-level forms regardless of whether their progress is monitored with off-grade-level forms. In addition, when a student is demonstrating progress on off-grade-level forms, we advise occasionally administering an on-grade-level progress monitoring form every 4 to 6 weeks. Once a student meets the end of year benchmark goal for the off-grade level with which they are being progress monitored, the student should be moved to on-grade-level progress monitoring.

©2026 University of Oregon. All rights reserved.

References

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

Al Otaiba, S., McDonald Connor, C., Foorman, B., Schatschneider, C., Greulich, L., & Sidler, J. F. (2009). Identifying and intervening with beginning readers who are at-risk for dyslexia: Advances in individualized classroom instruction. *Perspectives on language and literacy, 35*(4), 13–19. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4296731/

Alonzo, J., & Tindal, G. (2007). *The development of fifth-grade passage reading fluency measures for use in a progress monitoring assessment system* (Technical report No. 40). Behavioral Research and Teaching. http://www.brtprojects.org/publications/dl/54

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational Psychological Testing. (2014). Standards for educational and psychological testing. American Educational Research Association.

Baker, D. L., Biancarosa, G., Park, B. J., Bousselot, T., Smith, J. L. M., Baker, S. K., . . . Tindal, G. (2015). Validity of CBM measures of oral reading fluency and reading comprehension on high-stakes reading assessments in grades 7 and 8. *Reading and Writing, 28*, 57–104. https://doi.org/10.1007/s11145-014-9505-4

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445–459. https://doi.org/10.3758/bf03193014

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Biemiller, A. (2008). Words worth teaching. SRA/McGraw-Hill.

Biemiller, A. (2010). Words worth teaching: Closing the vocabulary gap. McGraw-Hill SRA.

Bizzocchi, A. (2017). How many phonemes does the English language have? *International Journal on Studies in English Language and Literature, 5*, 36–46. https://doi.org/10.20431/2347-3134.0510006

©2026 University of Oregon. All rights reserved.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Brysbaert, M., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Methods, 49*(4), 1520–1523. https://doi.org/10.3758/s13428-016-0811-4

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*, 261–304. https://doi.org/10.1177/0049124104268644

Burns, M. K., Ardoin, S. P., Parker, D. C., Hodgson, J., Klingbeil, D. A., & Scholin, S. E. (2009). Interspersal technique and behavioral momentum for reading word lists. *School Psychology Review, 38*, 428–434. https://doi.org/10.1080/02796015.2009.12087825

Catts, H. W., Petscher, Y., Schatschneider, C., Sittner Bridges, M., & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities, 42*, 163–176. https://doi.org/10.1177/0022219408326219

Chall, J. S. (1996). Stages of reading development (2nd ed.). Harcourt Brace.

Christ, T. J. (2006). Short-term estimates of growth using curriculum-based measurement of oral reading fluency: Estimating standard error of the slope to construct confidence intervals. *School Psychology Review, 35*, 128–133. https://doi.org/10.1080/02796015.2006.12088006

Christ, T. J., & Ardoin, S. P. (2009). Curriculum-based measurement of oral reading: Passage equivalence and probe-set development. *Journal of School Psychology, 47*, 55–75. https://doi.org/10.1016/j.jsp.2008.09.004

Christ, T. J., Silberglitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review, 39*, 447–462. https://doi.org/10.1080/02796015.2010.12087765

Christ, T. J., Zopluoglu, C., Long, J. D., & Monaghen, B. D. (2012). Curriculum-based measurement of

©2026 University of Oregon. All rights reserved.

oral reading: Quality of progress monitoring outcomes. *Exceptional Children, 78*, 356–373. https://www.proquest.com/scholarly-journals/curriculum-based-measurement-oral-reading-quality/docview/926833518/se-2

Clemens, N. H., Shapiro, E. S., & Thoemmes, F. (2011). Improving the efficacy of first grade reading screening: An investigation of word identification fluency with other early literacy indicators. *School Psychology Quarterly, 26*, 231–244. https://doi.org/10.1037/a0025173

Conoyer, S. J., Lembke, E. S., Hosp, J. L., Espin, C. A., Hosp, M. K., & Poch, A. L. (2017). Getting more from your maze: Examining differences in distractors. *Reading & Writing Quarterly, 33*(2), 141–154. https://doi.org/10.1080/10573569.2016.1142913

Cummings, K. D., Park, Y., & Bauer Schaper, H. A. (2013). Form effects on DIBELS next Oral Reading Fluency progress-monitoring passages. *Assessment for Effective Intervention, 38*, 91–104. https://doi.org/10.1177/1534508412447010

Dale, E., & O'Rourke, J. (1981). The living word vocabulary: A national vocabulary inventory. World Book-Childcraft International.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15*(3), 358–374. https://doi.org/10.1080/02796015.1986.12085239

Deno, S. L., Reschly, A. L., Lembke, E. S., Magnusson, D., Callender, S. A., Windram, H., & Stachel, N. (2009). Developing a school-wide progress-monitoring system. *Psychology in the Schools, 46*, 44–55. https://doi.org/10.1002/pits.20353

DiStefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*(1), 20. https://doi.org/10.7275/da8t-4g52

Duranovic, M., Senka, S., & Babic-Gavric, B. (2018). Influence of increased letter spacing and font type on the reading ability of dyslexic children. *Annals of Dyslexia, 68*(3), 218–228. https://doi.org/10.1007/s11881-018-0164-z

Dyson, M. C., & Haselgrove, M. (2001). The influence of reading speed and line length on the effectiveness of reading from screen. *International Journal of Human-Computer Studies,*

©2026 University of Oregon. All rights reserved.

54(4), 585–612. https://doi.org/10.1006/ijhc.2001.0458

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*, 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Espin, C. A., Wallace, T., Lembke, E., Campbell, H., & Long, J. D. (2010). Creating a progress-monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research & Practice, 25*, 60–75. https://doi.org/10.1111/j.1540-5826.2010.00304.x

Francis, D. J., Santi, K. L., Barr, C., Fletcher, J. M., Varisco, A., & Foorman, B. R. (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology, 46*, 315–342. https://doi.org/10.1016/j.jsp.2007.06.003

Fry, E. (2004). Phonics: A large phoneme - grapheme frequency count revised. *Journal of Literacy Research, 36*(1), 85–98. https://doi.org/10.1207/s15548430jlr3601_5

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

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention: A decade later. *Journal of Learning Disabilities, 45*, 195–203. https://doi.org/10.1177/0022219412442150

Gersten, R. M., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2009). *Assisting students struggling with reading: Response to intervention and multi-tier intervention for reading in the primary grades: A practice guide* (Report No. NCE 2009-4045). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/rti_reading_pg_021809.pdf

Geverdt, D. E., & National Center for Education Statistics. (2015). *Education demographic and geographic estimates program (edge): Locale boundaries user's manual* (NCES No. 2016-012). ERIC Clearinghouse, National Center for Education Statistics (ED). https://books.google.com/books?id=SWofvwEACAAJ

Good, R. H., & Kaminski, R. A. (2009). DIBELS next. Voyager Sopris Learning.

Habibzadeh, F., Habibzadeh, P., & Yadollahie, M. (2016). On determining the most appropriate test

©2026 University of Oregon. All rights reserved.

cut-off value: The case of tests with continuous results. *Biochemia medica, 26*(3), 297–307. https://doi.org/10.11613/BM.2016.034

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. https://doi.org/10.1080/10705519909540118

Hughes, L., & Wilkins, A. (2000). Typography in children's reading schemes may be suboptimal: Evidence from measures of reading rate. *Journal of Research in Reading, 23*(3), 314–324. https://doi.org/10.1111/1467-9817.00126

International Dyslexia Association. (2020). *Universal screening: K–2 reading* International Dyslexia Association. https://dyslexiaida.org/universal-screening-k-2-reading

January, S.-A. A., & Ardoin, S. P. (2012). The impact of context and word type on students' maze task accuracy. *School Psychology Review, 41*, 262–271. https://doi.org/10.1080/02796015.2012.12087508

January, S.-A. A., Ardoin, S. P., Christ, T. J., Eckert, T. L., & White, M. J. (2016). Evaluating the interpretations and use of curriculum-based measurement in reading and word lists for universal screening in first and second grade. *School Psychology Review, 45*(3), 310–326. https://doi.org/10.17105/SPR45-3.310-326

Jones, M. N., & Mewhort, D. J. K. (2004). Case-sensitive letter and bigram frequency counts from large-scale English corpora. *Behavior Research Methods, Instruments, & Computers, 36*(3), 388–396. https://doi.org/10.3758/bf03195586

Katzir, T., Hershko, S., & Halamish, V. (2013). The effect of font size on reading comprehension on second and fifth grade children: Bigger is not always better. *PloS one, 8*(9), e74061–e74061. https://doi.org/10.1371/journal.pone.0074061

Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge Univ Pr.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press.

Kuster, S. M., van Weerdenburg, M., Gompel, M., & Bosman, A. M. T. (2018). Dyslexie font does not benefit reading in children with or without dyslexia. *Annals of Dyslexia, 68*(1), 25-42. https://

©2026 University of Oregon. All rights reserved.

doi.org/10.1007/s11881-017-0154-6

Lockenvitz, S., Kuecker, K., & Ball, M. J. (2015). Evidence for the distinction between "consonantal-/r/" and "vocalic-/r/" in American English. *Clinical Linguistics & Phonetics, 29*(8-10), 613–622. https://doi.org/10.3109/02699206.2015.1047962

Lonigan, C. J., & Burgess, S. R. (2017). Dimensionality of reading skills with elementary-school-age children. *Scientific Studies of Reading, 21*(3), 239–253. https://doi.org/10.1080/10888438.2017.1285918

Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018). Examining the simple view of reading with elementary school children: Still simple after all these years. *Remedial and Special Education, 39*(5), 260–273. https://doi.org/10.1177/0741932518764833

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*(2), 203–208. https://doi.org/10.3758/BF03204766

Marinus, E., Mostard, M., Segers, E., Schubert, T. M., Madelaine, A., & Wheldall, K. (2016). A special font for people with dyslexia: Does it work and, if so, why? *Dyslexia, 22*(3), 233–244. https://doi.org/10.1002/dys.1527

Marston, D. B., & Magnusson, D. (1988). Curriculum-based measurement: District level implementation. In J. L. Graden, J. Zins, & M. E. Curtis (Eds.), *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 137–172). National Association of School Psychology.

McNamara, D., Graesser, A. C., Cai, Z., & Dai, J. (2013). *Coh-Metrix common core T.E.R.A. .* (Version 1.0). Science of Learning and Educational Technology (SoLET) Lab. https://soletlab.adaptiveliteracy.com:8443/

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). Automated evaluation of text and discourse with Coh-Metrix. Cambridge University Press https://doi.org/10.1017/CBO9780511894664

Meyer, M. S. (2000). The ability–achievement discrepancy: Does it contribute to an understanding of learning disabilities? *Educational Psychology Review, 12*(3), 315–337. https://doi.

©2026 University of Oregon. All rights reserved.

org/10.1023/A:1009070006373

National Institute of Child Health and Human Development. (2000). *Teaching children to read: An evidence-based assessment of scientific research literature on reading and its implications for reading instruction*. NICHD Clearinghouse. http://www.nichd.nih.gov/publications/pubs/nrp/documents/report.pdf

Nese, J. F., Biancarosa, G., Cummings, K. D., Kennedy, P. C., Alonzo, J., & Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across grades 1-8. *Journal of School Psychology, 51*, 625–642. https://doi.org/10.1016/j.jsp.2013.05.006

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

Norvig, P. (2012). *English letter frequency counts Mayzner revisited or ETAOIN SRHLDCU*. http://norvig.com/mayzner.html

O'Brien, B. A., Mansfield, J. S., & Legge, G. E. (2005). The effect of print size on reading speed in dyslexia. *Journal of Research in Reading, 28*(3), 332–349. https://doi.org/10.1111/j.1467-9817.2005.00273.x

Paris, S., & Hamilton, E. E. (2009). The development of children's reading comprehension. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 32–53). Lawrence Erllbaum.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*, 184–202. https://doi.org/10.1598/RRQ.40.2.3

Pepe, M. S. (2003). The statistical evaluation of medical tests for classification and prediction. Oxford https://doi.org/10.1093/oso/9780198509844.001.0001

Petscher, Y., & Kim, Y.-S. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *Journal of School Psychology, 49*, 107–129. https://doi.org/10.1016/j.jsp.2010.09.004

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

©2026 University of Oregon. All rights reserved.

Reed, D. K., Cummings, K. D., Schaper, A., Lynn, D., & Biancarosa, G. (2018). Accuracy in identifying students' miscues during oral reading: A taxonomy of scorers' mismarkings. *Reading and Writing, 32*(4), 1009–1035. https://doi.org/10.1007/s11145-018-9899-5

Riedel, B. W., & Samuels, S. J. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students [with commentary]. *Reading Research Quarterly, 42*, 546–567. http://www.jstor.org/stable/20068319

Santi, K. L., Barr, C., Khalaf, S., & Francis, D. J. (2016). Different approaches to equating oral reading fluency passages. In K. D. Cummings & Y. Petscher (Eds.), *The fluency construct: Curriculum-based measurement concepts and applications.* (pp. 223–265). Springer Science + Business Media. https://doi.org/10.1007/978-1-4939-2803-3_9

Shanahan, T., Kamil, M. L., & Tobin, A. W. (1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly, 17*, 229–255. http://www.jstor.org/stable/747485

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *Journal of Special Education, 34*, 164–172. https://doi.org/10.1177/002246690003400305

Shinn, M. R. (1989). Curriculum-based measurement: Assessing special children. Guilford.

Shinn, M. R., & Bamonto, S. (1998). Advanced applications of curriculum-based measurement: "Big ideas" and avoiding confusion. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 1–31). Guildford Press.

Smith, J. L. M., Cummings, K. D., Nese, J. F. T., Alonzo, J., Fien, H., & Baker, S. K. (2014). The relation of word reading fluency initial level and gains with reading outcomes. *School Psychology Review, 43*, 30–40. https://doi.org/10.1080/02796015.2014.12087452

Snow, C. E., Burns, M. S., & Griffin, P. (1998). Preventing reading difficulties in young children. National Academy Press. http://www.nap.edu/catalog/6023.html

Thurstone, L. L. (1935). The vectors of mind: Multiple-factor analysis for the isolation of primary traits. University of Chicago Press https://doi.org/10.1037/10018-000

U.S. Department of Commerce, Economics and Statistics Administration, & Bureau of the Census.

©2026 University of Oregon. All rights reserved.

(1994). *Geographic areas reference manual*. U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census. https://www.census.gov/programs-surveys/geography/guidance/geographic-areas-reference-manual.html

University of Iowa. (2015). Iowa assessments. Houghton Mifflin Harcourt.

Welch, C., & Dunbar, S. (2012). Iowa assessments. Houghton Mifflin.

Wery, J. J., & Diliberto, J. A. (2017). The effect of a specialized dyslexia font, opendyslexic, on reading rate and accuracy. *Annals of dyslexia, 67*(2), 114–127. https://doi.org/10.1007/s11881-016-0127-1

Wilkins, A., Cleave, R., Grayson, N., & Wilson, L. (2009). Typography for children may be inappropriately designed. *Journal of Research in Reading, 32*(4), 402–412. https://doi.org/10.1111/j.1467-9817.2009.01402.x

Wise Justin, C., Sevcik Rose, A., Morris Robin, D., Lovett Maureen, W., Wolf, M., Kuhn, M., . . . Schwanenflugel, P. (2010). The relationship between different measures of Oral Reading Fluency and reading comprehension in second-grade students who evidence different Oral Reading Fluency difficulties. *Language, Speech, and Hearing Services in Schools, 41*(3), 340–348. https://pubmed.ncbi.nlm.nih.gov/20421613/

©2026 University of Oregon. All rights reserved.

Appendix A: ORF Passage Statistics

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
1	Benchmark ORF.1.Beginning	Lucky Day	172	21	1.5	4-500L	95	72	76	62	42
	Benchmark ORF.1.Middle	Jack and Jill	190	22	1.2	4-500L	91	98	53	71	48
	Benchmark ORF.1.End	Our Pond	169	15	1.8	4-500L	98	49	93	94	44
	Progress Monitoring ORF.1.1	A Clean House	197	19	1.5	4-500L	99	52	66	99	78
	Progress Monitoring ORF.1.2	Brush Your Teeth	218	28	1.5	4-500L	59	98	89	98	54
	Progress Monitoring ORF.1.3	Hats	199	22	1.0	4-500L	57	95	90	99	50
	Progress Monitoring ORF.1.4	Glass Frogs	207	25	1.2	5-600L	71	97	65	41	38
	Progress Monitoring ORF.1.5	On the Bridge	172	19	1.5	4-500L	99	78	73	60	15
	Progress Monitoring ORF.1.6	Ants	197	23	0.6	5-600L	86	89	88	62	20
	Progress Monitoring ORF.1.7	The Yellow House	188	21	1.2	4-500L	87	77	50	56	21
	Progress Monitoring ORF.1.8	The Bus Ride	191	19	1.7	4-500L	96	74	61	76	78
	Progress Monitoring ORF.1.9	We Need Sleep	200	20	1.9	4-500L	93	83	56	94	46
	Progress Monitoring ORF.1.10	Breakfast in Bed	228	25	1.2	5-600L	92	93	89	45	97
	Progress Monitoring ORF.1.11	The Fishing Trip	205	22	0.7	4-500L	97	88	68	73	51
	Progress Monitoring ORF.1.12	Owls	221	29	1.3	4-500L	17	94	83	44	64
	Progress Monitoring ORF.1.13	Bears	180	17	1.7	4-500L	81	85	90	82	70
	Progress Monitoring ORF.1.14	A Walk with Mom	196	22	1.0	4-500L	97	92	85	70	94
	Progress Monitoring ORF.1.15	Cats and Dogs	238	26	1.6	4-500L	76	99	99	98	98
	Progress Monitoring ORF.1.16	My Bike	170	19	1.0	4-500L	99	86	66	84	28
	Progress Monitoring ORF.1.17	The Book Club	222	24	1.3	4-500L	94	91	59	62	92
	Progress Monitoring ORF.1.18	Hippos	218	25	1.9	4-500L	86	85	84	69	33
	Progress Monitoring ORF.1.19	The Video Game Fight	191	22	1.7	3-400L	98	96	76	92	88
	Progress Monitoring ORF.1.20	Not All Birds Can Fly	168	20	1.3	4-500L	89	97	67	82	93

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
2	Benchmark ORF.2.Beginning	Church Pears	225	20	2.4	4-500L	76	59	96	59	9
	Benchmark ORF.2.Middle	Pay Phones	201	17	2.9	4-500L	65	50	63	89	81
	Benchmark ORF.2.End	Puppy Love	201	17	2.7	4-500L	86	71	96	94	44
	Progress Monitoring ORF.2.1	Grandma and Grandpop	157	13	2.8	4-500L	98	52	90	81	28
	Progress Monitoring ORF.2.2	Old Peach Tree	225	23	2.8	5-600L	77	94	98	72	54
	Progress Monitoring ORF.2.3	The Plane	182	17	2.7	5-600L	44	92	92	57	98
	Progress Monitoring ORF.2.4	Firefighters	153	15	2.3	4-500L	86	89	99	98	81
	Progress Monitoring ORF.2.5	Nuts about Ice Cream	177	16	2.6	5-600L	94	72	98	99	31
	Progress Monitoring ORF.2.6	Farm Chores	226	23	2.7	5-600L	80	83	94	70	81
	Progress Monitoring ORF.2.7	An Island of Mangroves	220	27	2.6	4-500L	48	95	97	25	24
	Progress Monitoring ORF.2.8	Ice Cream Cones	225	24	2.8	5-600L	37	90	88	77	86
	Progress Monitoring ORF.2.9	The Lost Cat	218	25	2.9	6-700L	81	86	94	14	6
	Progress Monitoring ORF.2.10	Clean Your Room	239	27	2.4	4-500L	63	92	65	79	84
	Progress Monitoring ORF.2.11	Saving Money	222	23	2.8	4-500L	90	90	43	98	64
	Progress Monitoring ORF.2.12	Ten Legs	170	27	2.8	4-500L	76	93	36	41	16
	Progress Monitoring ORF.2.13	Dad's Helper	203	18	2.8	5-600L	98	69	45	71	92
	Progress Monitoring ORF.2.14	The Lion's Pride	173	17	3.7	5-600L	45	91	71	82	55
	Progress Monitoring ORF.2.15	Keeping Track	255	24	2.3	5-600L	65	93	87	94	86
	Progress Monitoring ORF.2.16	Jungle and Forest	205	26	2.7	4-500L	3	84	21	92	2
	Progress Monitoring ORF.2.17	My Aunt	207	17	2.6	5-600L	100	27	28	98	73
	Progress Monitoring ORF.2.18	Pigs	147	16	2.6	4-500L	75	89	68	92	58
	Progress Monitoring ORF.2.19	Washing Dishes	171	16	2.6	5-600L	81	84	85	82	81
	Progress Monitoring ORF.2.20	Snow Days	182	14	2.9	4-500L	97	62	95	89	99

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
3	Benchmark ORF.3.Beginning	Sponges	258	24	4.1	5-600L	57	91	97	53	90
	Benchmark ORF.3.Middle	Why We Need Water	202	19	3.6	4-500L	97	73	93	98	73
	Benchmark ORF.3.End	Trees	201	16	2.8	4-500L	67	76	98	98	58
	Progress Monitoring ORF.3.1	To Be a Poet	217	17	3.4	5-600L	93	42	97	60	33
	Progress Monitoring ORF.3.2	Clara's Secret Life	244	22	3.1	5-600L	68	58	99	77	19
	Progress Monitoring ORF.3.3	Mexican Food	238	26	3.9	6-700L	7	97	97	22	59
	Progress Monitoring ORF.3.4	The Chest	212	19	3.5	6-700L	68	79	98	26	14
	Progress Monitoring ORF.3.5	The Tight Rope	216	19	3.8	6-700L	51	91	99	30	50
	Progress Monitoring ORF.3.6	A Friend in Osaka	394	47	3.3	5-600L	86	90	94	18	11
	Progress Monitoring ORF.3.7	Meet Your Spleen	229	23	3.8	5-600L	44	75	27	37	53
	Progress Monitoring ORF.3.8	The Best Birthday Ever	223	22	3.3	4-500L	92	92	90	70	99
	Progress Monitoring ORF.3.9	The Man Who Lived in a Hollow Tree	209	18	3.8	6-700L	94	66	90	37	33
	Progress Monitoring ORF.3.10	Turkey Vultures	271	26	4.0	7-800L	67	79	82	38	74
	Progress Monitoring ORF.3.11	Let's Play	152	13	4.0	5-600L	94	95	91	89	80
	Progress Monitoring ORF.3.12	First Snow	233	20	3.5	6-700L	72	82	99	38	53
	Progress Monitoring ORF.3.13	The Alley Cat	270	21	3.1	5-600L	72	62	99	65	34
	Progress Monitoring ORF.3.14	The Trolley	266	25	3.6	5-600L	90	78	93	25	43
	Progress Monitoring ORF.3.15	The Cat's Meow	206	17	5.1	4-500L	87	56	89	97	83
	Progress Monitoring ORF.3.16	Shadow Puppet Theater	267	23	4.0	5-600L	69	88	96	62	95
	Progress Monitoring ORF.3.17	At the Pond	254	20	3.5	4-500L	82	39	96	42	11
	Progress Monitoring ORF.3.18	The Baker's Daughter	177	18	3.1	4-500L	97	88	85	98	97
	Progress Monitoring ORF.3.19	Recycling	213	20	4.0	5-600L	58	96	61	40	92
	Progress Monitoring ORF.3.20	A New Room for Ruby and Ron	275	25	3.3	4-500L	85	81	93	79	62

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
4	Benchmark ORF.4.Beginning	The Raft	221	18	4.0	6-700L	49	78	98	27	14
	Benchmark ORF.4.Middle	Honesty	292	21	4.2	5-600L	99	42	75	95	64
	Benchmark ORF.4.End	Sunset at the Beach	250	20	5.0	6-700L	38	85	99	12	83
	Progress Monitoring ORF.4.1	Juan's Animals	252	22	4.3	5-600L	60	73	98	25	51
	Progress Monitoring ORF.4.2	A Nest Made of Paper	228	19	5.0	6-700L	33	62	96	26	18
	Progress Monitoring ORF.4.3	Storm at Sea	271	22	3.9	6-700L	43	83	97	6	69
	Progress Monitoring ORF.4.4	Ice Cubes	271	23	4.3	5-600L	59	90	98	98	91
	Progress Monitoring ORF.4.5	Deborah and the Sunset	217	19	4.5	5-600L	72	79	87	26	48
	Progress Monitoring ORF.4.6	Earthworms Are the Best	178	15	5.1	610-800L	41	95	95	29	93
	Progress Monitoring ORF.4.7	The Great Inventor	253	23	4.5	5-600L	72	93	81	36	97
	Progress Monitoring ORF.4.8	Black Cats	241	22	4.2	6-700L	45	69	65	86	72
	Progress Monitoring ORF.4.9	Jewelry	198	14	4.9	5-600L	97	45	67	97	99
	Progress Monitoring ORF.4.10	My Sister and I	327	26	4.9	8-900L	93	88	96	29	77
	Progress Monitoring ORF.4.11	Islands	274	28	4.9	6-700L	22	87	88	61	57
	Progress Monitoring ORF.4.12	The Runner	265	24	5.2	7-800L	79	78	63	27	50
	Progress Monitoring ORF.4.13	Build an Obstacle Course	271	24	4.6	7-800L	50	87	45	61	96
	Progress Monitoring ORF.4.14	The River Crossing	221	17	4.8	6-700L	61	60	98	48	22
	Progress Monitoring ORF.4.15	Living in Space	261	18	5.0	6-700L	83	68	95	76	97
	Progress Monitoring ORF.4.16	Skating	278	24	4.0	6-700L	95	87	84	82	55
	Progress Monitoring ORF.4.17	To Spit or Not to Spit	207	19	5.0	7-800L	30	89	45	44	46
	Progress Monitoring ORF.4.18	Red Barns	262	23	4.5	8-900L	35	95	92	69	98
	Progress Monitoring ORF.4.19	Maple Taffy	274	21	4.9	7-800L	39	71	98	39	47
	Progress Monitoring ORF.4.20	Blackberry Picking	263	22	4.1	6-700L	35	81	99	4	56

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
5	Benchmark ORF.5.Beginning	Breathing	296	24	5.2	8-900L	63	79	81	91	61
	Benchmark ORF.5.Middle	Animal Tools	320	23	5.8	9-1000L	42	78	99	79	97
	Benchmark ORF.5.End	The Project Next Door	286	18	5.9	9-1000L	79	47	99	51	46
	Progress Monitoring ORF.5.1	Sam the Cat	319	22	5.9	9-1000L	88	24	96	31	42
	Progress Monitoring ORF.5.2	Koda's Big Day	256	16	5.3	6-700L	54	42	99	42	14
	Progress Monitoring ORF.5.3	The North American Beaver	319	23	6.1	9-1000L	40	71	97	73	98
	Progress Monitoring ORF.5.4	Madelines	253	15	5.4	7-800L	95	9	99	94	27
	Progress Monitoring ORF.5.5	Mom, the Pastry Chef	295	21	5.4	8-900L	77	43	53	12	75
	Progress Monitoring ORF.5.6	Underneath the Mistletoe	261	23	5.7	8-900L	33	88	95	25	91
	Progress Monitoring ORF.5.7	Annie and the Lady	315	21	5.8	8-900L	95	55	92	82	70
	Progress Monitoring ORF.5.8	Glow Worms	292	19	6.0	10-1100L	14	40	84	25	10
	Progress Monitoring ORF.5.9	Eating Contest	314	21	5.7	7-800L	85	37	98	49	17
	Progress Monitoring ORF.5.10	The Barge	256	18	5.3	9-1000L	32	58	99	24	13
	Progress Monitoring ORF.5.11	Not A Duck	246	20	5.5	8-900L	64	71	76	33	89
	Progress Monitoring ORF.5.12	Milking the Cows	254	16	5.6	6-700L	47	45	99	39	14
	Progress Monitoring ORF.5.13	It's All a Mystery	252	17	5.8	8-900L	93	37	76	79	86
	Progress Monitoring ORF.5.14	Teaching Beatrix	259	25	5.5	5-600L	79	80	87	43	23
	Progress Monitoring ORF.5.15	The Story of Tea	332	31	5.8	7-800L	18	78	80	45	70
	Progress Monitoring ORF.5.16	Trains	283	21	5.4	9-1000L	18	86	95	80	91
	Progress Monitoring ORF.5.17	A Ball of Clay	325	17	5.7	9-1000L	96	24	93	93	92
	Progress Monitoring ORF.5.18	Up and Down	272	16	5.8	8-900L	63	49	99	23	10
	Progress Monitoring ORF.5.19	Scootering	317	21	5.7	8-900L	85	44	94	50	29
	Progress Monitoring ORF.5.20	Swimming	279	18	5.6	7-800L	90	56	98	77	43

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
6	Benchmark ORF.6.Beginning	Helen Keller	299	21	6.5	8-900L	72	76	79	68	72
	Benchmark ORF.6.Middle	Sloths and Monkeys	257	19	6.5	9-1000L	46	77	87	95	92
	Benchmark ORF.6.End	Changing Bedtime	324	23	6.4	8-900L	74	73	34	22	95
	Progress Monitoring ORF.6.1	Coming of Age	268	21	6.9	9-1000L	17	71	74	12	85
	Progress Monitoring ORF.6.2	The Wise and Strong Monkeys	332	21	5.7	9-1000L	73	58	83	63	86
	Progress Monitoring ORF.6.3	Fiona's Bad Day	373	20	6.8	9-1000L	96	27	85	77	89
	Progress Monitoring ORF.6.4	Hard Work	284	22	6.3	8-900L	52	90	23	50	96
	Progress Monitoring ORF.6.5	Forest Fires	286	19	6.7	9-1000L	28	50	99	48	99
	Progress Monitoring ORF.6.6	Talking Trees	311	19	6.6	8-900L	93	29	97	67	62
	Progress Monitoring ORF.6.7	Animal Behavior During an Eclipse	281	24	6.6	8-900L	44	80	69	2	68
	Progress Monitoring ORF.6.8	Thomas Young and Light Waves	303	19	7.2	10-1100L	44	64	98	55	51
	Progress Monitoring ORF.6.9	Still a Mystery to Me	327	19	6.2	9-1000L	94	39	93	63	72
	Progress Monitoring ORF.6.10	Animal Minds	325	24	7.1	8-900L	75	76	58	45	97
	Progress Monitoring ORF.6.11	Soap	322	25	6.4	9-1000L	55	74	94	87	86
	Progress Monitoring ORF.6.12	Ella Fitzgerald	269	19	6.9	9-1000L	78	51	92	61	53
	Progress Monitoring ORF.6.13	Lego	312	22	7.3	9-1000L	49	73	90	17	91
	Progress Monitoring ORF.6.14	The Umbrella	259	22	6.6	8-900L	7	94	93	36	96
	Progress Monitoring ORF.6.15	Control Control	317	23	6.8	8-900L	72	69	46	50	94
	Progress Monitoring ORF.6.16	Tower Beach	320	19	6.6	10-1100L	93	59	95	34	95
	Progress Monitoring ORF.6.17	Code Braille	258	16	6.4	9-1000L	50	67	70	53	57
	Progress Monitoring ORF.6.18	Apple-Picking Time	314	18	6.0	8-900L	91	31	95	69	44
	Progress Monitoring ORF.6.19	Drying Clothing	322	31	5.6	7-800L	35	93	29	46	71
	Progress Monitoring ORF.6.20	Nora's Town	298	18	6.9	8-900L	42	61	95	15	87

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
7	Benchmark ORF.7.Beginning	Coyotes and Wolves	279	20	7.4	10-1100L	25	60	98	83	68
	Benchmark ORF.7.Middle	Fizzy Water	361	29	7.4	9-1000L	27	82	85	77	88
	Benchmark ORF.7.End	Prize Winning Vegetables	370	22	7.8	10-1100L	53	19	87	52	55
	Progress Monitoring ORF.7.1	Zombees	367	24	7.5	10-1100L	27	73	88	67	94
	Progress Monitoring ORF.7.2	Captain Marshmallow	268	18	7.3	8-900L	74	7	76	29	96
	Progress Monitoring ORF.7.3	On Top of the Ferris Wheel	374	19	7.8	11-1200L	90	40	87	50	68
	Progress Monitoring ORF.7.4	Benjamin's Alarm Clock	366	26	7.6	9-1000L	90	37	95	42	26
	Progress Monitoring ORF.7.5	The Standing Invitation	373	17	8.5	11-1200L	87	19	99	57	75
	Progress Monitoring ORF.7.6	Two People in a Boat	325	17	7.3	11-1200L	67	46	92	18	95
	Progress Monitoring ORF.7.7	Dirty Rivers	368	32	7.8	7-800L	14	91	83	44	79
	Progress Monitoring ORF.7.8	Stopover in Tornado Alley	333	22	7.0	9-1000L	77	68	79	6	72
	Progress Monitoring ORF.7.9	Waiting for the Easter Bunny	356	21	6.8	10-1100L	85	18	83	31	23
	Progress Monitoring ORF.7.10	Four-Year-Old Peach Pie	369	17	8.8	11-1200L	87	22	99	56	95
	Progress Monitoring ORF.7.11	Taming a Wild Horse	300	17	6.9	9-1000L	50	39	98	62	65
	Progress Monitoring ORF.7.12	Glass Flowers	379	25	7.1	9-1000L	69	57	98	40	95
	Progress Monitoring ORF.7.13	What is Capitalism?	329	25	7.7	8-900L	37	92	61	55	95
	Progress Monitoring ORF.7.14	Lucky Pennies	346	21	7.0	9-1000L	77	39	77	71	88
	Progress Monitoring ORF.7.15	The Lost Rocket	355	17	8.7	12-1300L	86	42	88	20	77
	Progress Monitoring ORF.7.16	American Blue Jeans	281	18	7.2	10-1100L	44	72	74	24	99
	Progress Monitoring ORF.7.17	The Ring	323	26	5.8	9-1000L	81	55	96	56	37
	Progress Monitoring ORF.7.18	Squirrel Life	371	20	7.8	10-1100L	93	25	77	74	70
	Progress Monitoring ORF.7.19	Battling Stings	360	23	7.1	10-1100L	35	70	74	28	98
	Progress Monitoring ORF.7.20	A Summer Sail	368	15	9.6	11-1200L	95	21	91	37	94

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
8	Benchmark ORF.8.Beginning	Crows	353	20	8.5	11-1200L	38	55	93	20	53
	Benchmark ORF.8.Middle	Government	368	29	8.7	8-900L	11	86	31	28	70
	Benchmark ORF.8.End	Digital Music Recording	371	25	8.8	9-1000L	25	91	72	59	99
	Progress Monitoring ORF.8.1	Jellyfish	317	23	7.9	9-1000L	25	69	84	7	49
	Progress Monitoring ORF.8.2	My Fantastic Memory	369	20	8.2	9-1000L	97	37	79	83	62
	Progress Monitoring ORF.8.3	New Eyeglasses	397	22	8.0	10-1100L	65	56	97	22	90
	Progress Monitoring ORF.8.4	Elephant, Giraffe, and Kangaroo	369	28	8.6	9-1000L	31	95	47	52	99
	Progress Monitoring ORF.8.5	Margaret Mead	341	26	9.2	9-1000L	20	82	19	18	53
	Progress Monitoring ORF.8.6	The Dynamics of Snow	306	22	9.1	9-1000L	26	30	37	6	8
	Progress Monitoring ORF.8.7	How I Became a Bookworm	376	22	8.0	10-1100L	86	20	81	64	80
	Progress Monitoring ORF.8.8	The Earth When I Was a Child	321	18	7.6	9-1000L	36	52	90	2	26
	Progress Monitoring ORF.8.9	Sonar	345	23	9.1	10-1100L	30	64	85	93	61
	Progress Monitoring ORF.8.10	Commercial Shipping	355	25	8.4	10-1100L	31	87	92	64	88
	Progress Monitoring ORF.8.11	The Speech That Was Never Given	348	18	9.0	10-1100L	42	69	46	51	93
	Progress Monitoring ORF.8.12	Colorful Language	343	22	8.8	10-1100L	22	77	46	37	76
	Progress Monitoring ORF.8.13	Pizza Night	336	18	9.0	11-1200L	36	37	99	54	70
	Progress Monitoring ORF.8.14	The Golden Age of Radio	363	23	8.5	9-1000L	23	68	93	37	78
	Progress Monitoring ORF.8.15	Three Famous Authors	348	25	9.0	9-1000L	36	74	42	20	77
	Progress Monitoring ORF.8.16	Big Money	346	19	8.4	11-1200L	16	55	84	22	81
	Progress Monitoring ORF.8.17	Bats	363	23	8.8	10-1100L	52	41	67	31	57
	Progress Monitoring ORF.8.18	The Giant Duck	342	21	8.1	10-1100L	34	40	91	8	60
	Progress Monitoring ORF.8.19	Ancient Little Farmers	375	22	9.1	11-1200L	24	55	88	46	61
	Progress Monitoring ORF.8.20	Time Travel	348	25	9.0	9-1000L	36	74	42	20	77

Note. W = word count. S = sentence count. FKGL = Flesch-Kincaid Grade Level. SS=Syntactic Simplicity. WC=Word Concreteness. RC=Referential Cohesion. DC=Deep Cohesion. W, S, and FKGL were drawn from Microsoft Word. The last five columns are Coh-Metrix indices ranging from 1 to 99 where higher values indicate more of a characteristic and less text complexity. Higher values of narrativity indicates passages that are more story-like. Higher values of syntactic simplicity indicate passages that use simpler sentence structures more frequently, although sentences containing dialogue can sometimes inflate these values. Higher values of word concreteness indicate passages that include more concrete rather than abstract words. Higher values of referential cohesion indicate more overlap in words and propositions in a passage. Higher values of deep cohesion indicate more frequent connectives in passages.

©2026 University of Oregon. All rights reserved.

Appendix B: Maze Passage Statistics

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
2	Benchmark Maze.2.Beginning	Working Together	371	38	2.2	5-600L	94	83	81	67	79
	Benchmark Maze.2.Middle	The New Kid	401	38	2.5	5-600L	95	75	84	51	92
	Benchmark Maze.2.End	Summer Reading	403	46	2.6	5-600L	84	89	93	49	46
	Progress Monitoring Maze.2.1	The Chairs	356	36	3.4	5-600L	73	86	95	55	18
	Progress Monitoring Maze.2.2	The Airplane	366	44	2.6	5-600L	81	94	93	14	28
	Progress Monitoring Maze.2.3	The Starry Sky	372	42	2.6	5-600L	92	89	85	29	77
	Progress Monitoring Maze.2.4	Shadow Play	411	39	2.5	5-600L	84	69	82	80	69
	Progress Monitoring Maze.2.5	The Dishwasher	357	43	3	4-500L	91	97	78	55	51
	Progress Monitoring Maze.2.6	Tin	359	32	2.7	5-600L	79	95	78	71	99
	Progress Monitoring Maze.2.7	Staying with Grams and Pappy	416	43	2.3	410-600L	83	87	82	8	52
	Progress Monitoring Maze.2.8	Video Chat	401	42	2.8	410-600L	92	81	90	54	94
	Progress Monitoring Maze.2.9	Fishing Poles	403	38	2.3	410-600L	85	87	99	64	77
	Progress Monitoring Maze.2.10	One Spooky Night	425	48	2.0	410-600L	81	88	90	20	73

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
3	Benchmark Maze.3.Beginning	Brush Hogging	380	30	4.1	5-600L	73	51	98	50	32
	Benchmark Maze.3.Middle	The Secret Desert	414	35	3.7	5-600L	90	54	75	65	79
	Benchmark Maze.3.End	On the Trail	420	34	3.7	6-700L	87	84	77	40	57
	Progress Monitoring Maze.3.1	The Time Capsule	366	36	3.6	5-600L	83	88	62	23	29
	Progress Monitoring Maze.3.2	Grandpa's Snakes	360	28	3.9	5-600L	87	69	97	70	72
	Progress Monitoring Maze.3.3	Dear Pen Pal	419	37	4.3	6-700L	61	72	85	13	44
	Progress Monitoring Maze.3.4	Dandelion Salad	414	41	3.4	5-600L	82	85	85	40	62
	Progress Monitoring Maze.3.5	Brain Freeze	360	32	3.8	5-600L	73	92	75	37	99
	Progress Monitoring Maze.3.6	The West	407	33	4	6-700L	79	68	93	38	76
	Progress Monitoring Maze.3.7	The Dinner Party	398	34	3.9	410-600L	77	81	90	51	34
	Progress Monitoring Maze.3.8	The Record	425	43	3.7	610-800L	82	79	80	23	31
	Progress Monitoring Maze.3.9	The Essay	422	33	4.4	610-800L	93	59	70	31	59
	Progress Monitoring Maze.3.10	Atlas	397	37	3.8	610-800L	70	72	75	36	62

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
4	Benchmark Maze.4.Beginning	Working on Cars	410	27	5	8-900L	85	58	97	30	37
	Benchmark Maze.4.Middle	Lucie's Snow	457	30	5.6	8-900L	79	48	99	48	28
	Benchmark Maze.4.End	The Hill	497	36	5.7	8-900L	82	47	83	55	40
	Progress Monitoring Maze.4.1	The Kellogg Brothers	413	38	4.9	7-800L	40	89	88	33	86
	Progress Monitoring Maze.4.2	Two Little Monkeys	378	29	5.5	5-600L	51	88	99	78	50
	Progress Monitoring Maze.4.3	Tom's Supper	405	28	4.8	7-800L	57	52	99	40	27
	Progress Monitoring Maze.4.4	Homemade Ice Cream	414	31	4.9	6-700L	53	87	99	66	62
	Progress Monitoring Maze.4.5	Maya and the Manatees	398	33	5.3	7-800L	78	71	90	38	31
	Progress Monitoring Maze.4.6	Butterflies and Moths	426	34	5.3	9-1000L	34	84	69	77	66
	Progress Monitoring Maze.4.7	Getting Back on the Bike	351	23	4.9	810-1000L	92	46	97	71	46
	Progress Monitoring Maze.4.8	Another Ice Cream Headache	436	36	5.2	810-1000L	67	73	72	15	45
	Progress Monitoring Maze.4.9	Stargazing at the Lakehouse	422	29	4.8	810-1000L	91	67	95	37	74
	Progress Monitoring Maze.4.10	Nesting Birds	483	34	5.0	810-1000L	41	81	98	39	77

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
5	Benchmark Maze.5.Beginning	The iPhone	486	32	6.9	9-1000L	58	70	76	21	86
	Benchmark Maze.5.Middle	New Slang	507	37	6.9	8-900L	52	82	20	25	85
	Benchmark Maze.5.End	Languages	483	40	7.1	8-900L	48	8	38	93	63
	Progress Monitoring Maze.5.1	A Boy Named Fridge	457	31	5.5	8-900L	96	48	92	57	78
	Progress Monitoring Maze.5.2	Lost and Found	498	33	6.1	9-1000L	79	57	99	64	59
	Progress Monitoring Maze.5.3	Basketball	484	28	7	10-1100L	62	32	73	87	74
	Progress Monitoring Maze.5.4	Versions of a Folk Tale	490	31	6.8	9-1000L	38	58	98	48	59
	Progress Monitoring Maze.5.5	A Memory Palace	467	30	6.7	9-1000L	71	44	92	73	85
	Progress Monitoring Maze.5.6	Liam the Sea Creature	457	26	6.9	9-1000L	89	35	89	67	40
	Progress Monitoring Maze.5.7	Types of Tears	470	37	6.0	810-1000L	42	71	67	34	53
	Progress Monitoring Maze.5.8	In the Woods	459	26	5.9	1010-1200L	57	53	99	17	25
	Progress Monitoring Maze.5.9	Glacier Mice	459	27	6.1	810-1000L	61	62	87	40	74
	Progress Monitoring Maze.5.10	The Blizzard	553	35	5.8	810-1000L	77	55	96	31	90

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
6	Benchmark Maze.6.Beginning	How to Make a Woodcut	444	26	7	10-1100L	36	68	73	75	85
	Benchmark Maze.6.Middle	Blizzards of the United States	525	32	6.6	10-1100L	27	79	91	14	95
	Benchmark Maze.6.End	Popcorn	512	34	7.8	10-1100L	23	76	90	33	96
	Progress Monitoring Maze.6.1	The Portal	475	31	6.5	9-1000L	66	53	41	23	25
	Progress Monitoring Maze.6.2	Sacred Saguaro	478	32	7.4	10-1100L	38	68	89	27	97
	Progress Monitoring Maze.6.3	Unintended Consequences	462	34	7.7	9-1000L	39	89	34	21	89
	Progress Monitoring Maze.6.4	Sea Stars	478	30	7.2	10-1100L	42	56	96	40	45
	Progress Monitoring Maze.6.5	Robert McCloskey	495	26	8.2	11-1200L	63	50	98	37	99
	Progress Monitoring Maze.6.6	Kimble	480	26	7.1	10-1100L	94	24	86	77	77
	Progress Monitoring Maze.6.7	The Great Library of Alexandria	483	35	7.6	810-1000L	25	74	68	32	72
	Progress Monitoring Maze.6.8	Are Hedgehogs and Porcupines Related?	458	34	6.4	610-800L	46	72	82	51	67
	Progress Monitoring Maze.6.9	The First Day of Summer	506	27	7.3	1010-1200L	81	26	94	35	60
	Progress Monitoring Maze.6.10	Simone Biles	452	32	7.2	810-1000L	75	59	78	57	83

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
7	Benchmark Maze.7.Beginning	Fireflies	453	32	8.1	9-1000L	26	78	79	14	78
	Benchmark Maze.7.Middle	Archimedes	451	30	8.9	9-1000L	52	54	84	58	40
	Benchmark Maze.7.End	The Day the Sky Turned Black	500	25	8.2	10-1100L	66	51	96	53	71
	Progress Monitoring Maze.7.1	How to Win an Argument	549	37	8.6	10-1100L	73	68	7	54	96
	Progress Monitoring Maze.7.2	Metal Eating Plants	495	34	7.6	9-1000L	21	70	90	49	80
	Progress Monitoring Maze.7.3	Why Don't School Buses Have Seatbelts?	457	23	9.4	12-1300L	29	43	91	77	28
	Progress Monitoring Maze.7.4	Wilma Rudolph	475	25	8.6	10-1100L	91	20	90	85	72
	Progress Monitoring Maze.7.5	The Sun	481	28	8.2	10-1100L	43	28	91	75	79
	Progress Monitoring Maze.7.6	Japanese Flower Arrangement	464	31	8.5	9-1000L	16	62	80	19	71
	Progress Monitoring Maze.7.7	The Cicadas are Coming!	514	34	7.7	1010-1200L	26	73	87	8	81
	Progress Monitoring Maze.7.8	Bamboo	524	36	7.1	810-1000L	21	92	93	44	100
	Progress Monitoring Maze.7.9	Diving Deep	512	31	7.9	1010-1200L	54	53	94	34	76
	Progress Monitoring Maze.7.10	Metal Detector	501	27	8.1	1010-1200L	60	26	98	11	66

©2026 University of Oregon. All rights reserved.

G	Period	Title	W	S	FKGL	Lexile	Narrativity	SS	WC	RC	DC
8	Benchmark Maze.8.Beginning	Backstrap Weaving	456	26	8	10-1100L	33	61	84	49	88
	Benchmark Maze.8.Middle	The Circus	538	32	8.6	10-1100L	42	40	95	71	85
	Benchmark Maze.8.End	Space Shuttles	533	35	8.8	10-1100L	15	87	77	34	76
	Progress Monitoring Maze.8.1	The Ukulele	494	33	8.8	9-1000L	18	72	73	46	84
	Progress Monitoring Maze.8.2	The Perseids	442	22	9.6	11-1200L	31	40	88	60	43
	Progress Monitoring Maze.8.3	Opera Singers	492	24	10	11-1200L	43	35	72	64	10
	Progress Monitoring Maze.8.4	Hurricane	454	28	9.2	10-1100L	20	58	95	61	53
	Progress Monitoring Maze.8.5	What is a Victorian?	458	31	9.7	10-1100L	11	74	61	27	62
	Progress Monitoring Maze.8.6	Double Dutch	472	24	8.8	11-1200L	33	49	90	39	85
	Progress Monitoring Maze.8.7	A Hike to Remember	519	30	8.4	1010-1200L	39	42	99	9	20
	Progress Monitoring Maze.8.8	Swimming in the Ancient World	476	23	8.9	1010-1200L	21	59	97	10	82
	Progress Monitoring Maze.8.9	Deep Sea Treasure	520	35	8.7	810-1000L	21	56	65	3	34
	Progress Monitoring Maze.8.10	Backyard Foxes	512	26	8.7	1210-1400L	64	43	96	9	70

Note. W = word count. S = sentence count. FKGL = Flesch-Kincaid Grade Level. SS=Syntactic Simplicity. WC=Word Concreteness. RC=Referential Cohesion. DC=Deep Cohesion. W, S, and FKGL were drawn from Microsoft Word. The last five columns are Coh-Metrix indices ranging from 1 to 99 where higher values indicate more of a characteristic and less text complexity. Higher values of narrativity indicate passages that are more story-like. Higher values of syntactic simplicity indicate passages that use simpler sentence structures more frequently, although sentences containing dialogue can sometimes inflate these values. Higher values of word concreteness indicate passages that include more concrete rather than abstract words. Higher values of referential cohesion indicate more overlap in words and propositions in a passage. Higher values of deep cohesion indicate more frequent connectives in passages

©2026 University of Oregon. All rights reserved.

Appendix C: DIBELS 8th Edition Pronunciation Guide

Phoneme	Phoneme Example	Phoneme	Phoneme Example
/a/	b***a***d	/b/	***b***at
/e/	b***e***d	/d/	***d***ad
/i/	b***i***d	/f/	***f***at
/o/	c***o***d, l***aw***	/g/	***g***et
/u/	b***u***d, “a” in ***a***bout	/h/	***h***ot
/A/	b***ai***t	/j/	***j***am, e***dge***
/E/	b***ea***d	/k/	***c***an, ***k***it, pi***ck***
/I/	t***ie***	/l/	***l***ap
/O/	b***oa***t	/m/	***m***an
/oo/	f***oo***d	/n/	***n***ap
/uu/	g***oo***d	/p/	***p***en
/ow/	c***ow***	/r/	***r***at, ***wr***ite
/oy/	p***oi***nt, b***oy***	/s/	***s***it, ***c***ity
/ar/	(1 phoneme) c***ar***	/t/	***t***ap
/air/	(1 phoneme) ch***air***	/v/	***v***an
/er/	(1 phoneme) h***er***, b***ir***d	/w/	***w***et
/ear/	(1 phoneme) cl***ea***r	/y/	***y***ak
/or/	(1 phoneme) f***or***	/z/	***z***oo
/oor/	(1 phoneme) p***our***	/ch/	***ch***in
		/sh/	***sh***ed
		/SH/	mea***s***ure, bei***ge***
		/th/	***th***in
		/TH/	***th***en
		/ng/	si***ng***

Note. Both voiced and unvoiced forms of ‘th’ and ‘sh’ are acceptable for nonsense words containing these digraphs.

©2026 University of Oregon. All rights reserved.

Appendix D: Administration and Scoring Fidelity Checklists

Letter Naming Fluency Fidelity Checklist

Pass	Needs practice	
☐	☐	1. Holds clipboard and timer so student cannot see what is recorded.
☐	☐	2. Places the student copy in front of the student.
☐	☐	3. Performs standardized directions verbatim.
☐	☐	4. Starts timer after saying "***Begin***".
☐	☐	5. Follows along and marks the scoring book as the student responds.
☐	☐	6. Administers acceptable prompts (i.e., hesitation and letter sound) correctly and when appropriate.
☐	☐	7. Applies scoring rules consistently and correctly.
☐	☐	8. Applies the discontinue rule correctly, if appropriate.
☐	☐	9. At the end of 60 seconds, puts a bracket (]) after the last letter named and says "***Stop***".
☐	☐	10. Accurately determines and records the total number of correct letter names in 60 seconds. Score is within 2 points of the expert examiner.

©2026 University of Oregon. All rights reserved.

Phonemic Segmentation Fluency Fidelity Checklist

Pass	Needs practice	
☐	☐	1. Holds clipboard and timer so student cannot see what is recorded.
☐	☐	2. Performs standardized directions verbatim, including the correction procedure, if applicable.
☐	☐	3. Starts timer after presenting the first word.
☐	☐	4. Follows along and marks the scoring book as the student responds.
☐	☐	5. As soon as the student is finished saying the sounds in the current word, presents the next word promptly and clearly.
☐	☐	6. Administers acceptable prompts correctly and when appropriate.
☐	☐	7. Applies scoring rules consistently and correctly.
☐	☐	8. Applies the discontinue rule correctly, if appropriate.
☐	☐	9. Stops at the end of 60 seconds and puts a bracket (]) after the last response.
☐	☐	10. Accurately determines and records the total number of correctly produced phonemes in 60 seconds. Score is within 2 points of the expert examiner.

©2026 University of Oregon. All rights reserved.

Nonsense Word Fluency Fidelity Checklist

Pass	Needs practice	
☐	☐	1. Holds clipboard and timer so student cannot see what is recorded.
☐	☐	2. Places student copy in front of the student.
☐	☐	3. Performs standardized directions verbatim, including the correction procedure when appropriate.
☐	☐	4. Starts timer after saying "***Begin***".
☐	☐	5. Follows along and marks the scoring book as the student responds.
☐	☐	6. Administers acceptable prompts correctly, if appropriate.
☐	☐	7. Applies scoring rules consistently and correctly.
☐	☐	8. Applies the discontinue rule correctly, if appropriate.
☐	☐	9. At the end of 60 seconds, puts a bracket (]) after the last sound provided and says "***Stop***".
☐	☐	10. Accurately determines and records the correct letter sounds produced and words read correctly within 60 seconds. Score is within 2 points of the expert examiner.

©2026 University of Oregon. All rights reserved.

Word Reading Fluency Fidelity Checklist

Pass	Needs practice	
☐	☐	1. Holds clipboard and timer so student cannot see what is recorded.
☐	☐	2. Places student copy in front of the student.
☐	☐	3. Performs standardized directions verbatim.
☐	☐	4. Starts timer when the student says the first word.
☐	☐	5. Follows along and marks the scoring book as the student responds.
☐	☐	6. Administers acceptable prompts correctly, if appropriate.
☐	☐	7. Applies scoring rules consistently and correctly.
☐	☐	8. Applies the discontinue rule correctly and when appropriate.
☐	☐	9. At the end of 60 seconds, puts a bracket (]) after the last word provided and says "***Stop***".
☐	☐	10. Accurately determines and records the number of words read correctly. Score is within 2 points of the expert examiner.

©2026 University of Oregon. All rights reserved.

Oral Reading Fluency Fidelity Checklist

Pass	Needs practice	
☐	☐	1. Holds clipboard and timer so student cannot see what is recorded.
☐	☐	2. Places student copy in front of the student.
☐	☐	3. Performs standardized directions verbatim.
☐	☐	4. Starts timer when the student says the first word.
☐	☐	5. Follows along and marks the scoring book as the student responds.
☐	☐	6. Administers acceptable prompts correctly, if appropriate.
☐	☐	7. Applies scoring rules consistently and correctly.
☐	☐	8. Applies the discontinue rule correctly and when appropriate.
☐	☐	9. At the end of 60 seconds, puts a bracket (]) after the last word read and says "***Stop***".
☐	☐	10. Accurately determines and records the number of words read correctly. Score is within 2 points of the expert examiner.

©2026 University of Oregon. All rights reserved.

Maze Fidelity Checklist

Pass	Needs practice	
☐	☐	1. Ensures each student has a copy of the Maze student materials, and students have written their name on it.
☐	☐	2. Performs standardized directions verbatim.
☐	☐	3. Starts timer after saying "***Begin***".
☐	☐	4. Administers acceptable prompts correctly, if appropriate.
☐	☐	5. At the end of 3 minutes, says "***Stop. Put your pencils down***."
☐	☐	6. Applies scoring rules consistently and correctly, using the scoring key.
☐	☐	7. Accurately determines and records the number of items answered correctly and incorrectly. Score is within 2 points of the expert examiner.
☐	☐	8. If not using the DIBELS Data System, accurately calculates the Maze Adjusted Score using the formula Correct – (Incorrect / 2).

©2026 University of Oregon. All rights reserved.

Appendix E: DIBELS 8th Edition Benchmark Goals

Key

Bold Numbers = The minimum score needed for ***core support***.

Legend

Support Needed	Risk Descriptor
Blue Goal = Core Support	*Negligible Risk* (nearly all students in this range score at or above the 40th percentile rank on criterion measure)
Green Range = Core Support	*Minimal Risk* **Bold** = the minimum score needed for core support (about 80% of students who score at or above the 40th percentile rank on criterion measure fall in this range or above)
Yellow Range = Strategic Support	*Some Risk* (about 80% of students who score below the 40th percentile on criterion measure fall in this range or below)
Red Range = Intensive Support	*At Risk* (about 80% of students who score below the 20th percentile on criterion measure fall in this range)

©2026 University of Oregon. All rights reserved.

Composite Score

Kindergarten to Third Grade

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Blue – Core Support	332+	393+	450+	354+	424+	480+	361+	423+	474+	365+	427+	467+
Green – Core Support	**306** - 331	**371** - 392	**420** - 449	**330** - 353	**389** - 423	**441** - 479	**329** - 360	**389** - 422	**439** - 473	**332** - 364	**393** - 426	**442** - 466
Yellow – Strategic Support	280 - 305	356 - 370	406 - 419	321 - 329	377 - 388	427 - 440	316 - 328	373 - 388	421 - 438	314 - 331	377 - 392	424 - 441
Red – Intensive Support	243 - 279	318 - 355	352 - 405	300 - 320	340 - 376	380 - 426	286 - 315	326 - 372	366 - 420	268 - 313	308 - 376	348 - 423

Fourth Grade to Eighth Grade

Support	Fourth Beg	Fourth Mid	Fourth End	Fifth Beg	Fifth Mid	Fifth End	Sixth Beg	Sixth Mid	Sixth End	Seventh Beg	Seventh Mid	Seventh End	Eighth Beg	Eighth Mid	Eighth End
Blue Core Support	368+	431+	461+	370+	421+	469+	364+	411+	454+	358+	407+	450+	378+	434+	478+
Green Core Support	**331** - 367	**399** - 430	**442** - 460	**335** - 369	**394** - 420	**449** - 468	**336** - 363	**386** - 410	**435** - 453	**336** - 357	**385** - 406	**430** - 449	**361** - 377	**404** - 433	**452** - 477
Yellow Strategic Support	310 - 330	380 - 398	421 - 441	313 - 334	380 - 393	436 - 448	313 - 335	370 - 385	419 - 434	315 - 335	374 - 384	417 - 429	345 - 360	391 - 403	437 - 451
Red Intensive Support	257 - 309	297 - 379	337 - 420	234 - 312	274 - 379	314 - 435	216 - 312	256 - 369	296 - 418	228 - 314	268 - 373	308 - 416	232 - 344	372 - 390	312 - 436

Letter Naming Fluency (LNF)

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Green – Core Support	**25+**	**37+**	**42+**	**42+**	**57+**	**59+**						
Yellow – Strategic Support	16 - 24	31 - 36	35 - 41	32 - 41	51 - 56	53 - 58						
Red – Intensive Support	0 - 15	0 - 30	0 - 34	0 - 31	0 - 50	0 - 52						

©2026 University of Oregon. All rights reserved.

Phoneme Segmentation Fluency (PSF)

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Blue – Core Support	15+	43+	53+	47+	57+	61+						
Green – Core Support	**5** - 14	**29** - 42	**44** - 52	**31** - 46	**43** - 56	**45** - 60						
Yellow – Strategic Support	1 - 4	23 - 28	37 - 43	19 - 30	34 - 42	37 - 44						
Red – Intensive Support	0	0 - 22	0 - 36	0 - 18	0 - 33	0 - 36						

Nonsense Word Fluency (NWF) – Correct Letter Sounds (CLS)

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Blue – Core Support	20+	36+	49+	47+	78+	87+	86+	103+	117+	121+	138+	141+
Green – Core Support	**9** - 19	**25** - 35	**31** - 48	**30** - 46	**52** - 77	**55** - 86	**50** - 85	**68** - 102	**76** - 116	**76** - 120	**94** - 137	**105** - 140
Yellow – Strategic Support	4 - 8	16 - 24	24 - 30	25 - 29	41 - 51	45 - 54	41 - 49	54 - 67	54 - 75	52 - 75	78 - 93	80 - 104
Red – Intensive Support	0 - 3	0 - 15	0 - 23	0 - 24	0 - 40	0 - 44	0 - 40	0 - 53	0 - 53	0 - 51	77 - 0	0 - 79

Nonsense Word Fluency (NWF) – Words Recoded Correctly (WRC)

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Blue – Core Support	--	9+	13+	16+	26+	28+	25+	36+	39+	34+	46+	45+
Green – Core Support	**1+**	**3** - 8	**7** - 12	**5** - 15	**14** - 25	**15** - 27	**15** - 24	**20** - 35	**22** - 38	**24** - 33	**30** - 45	**31** - 44
Yellow – Strategic Support	0	1 - 2	4 - 6	1 - 4	10 - 13	11 - 14	10 - 14	15 - 19	17 - 21	18 - 23	23 - 29	24 - 30
Red – Intensive Support	--	0	0 - 3	0	0 - 9	0 - 10	0 - 9	0 - 14	0 - 16	0 - 17	0 - 22	0 - 23

©2026 University of Oregon. All rights reserved.

Word Reading Fluency (WRF)

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Blue – Core Support	--	10+	18+	20+	33+	50+	50+	63+	70+	60+	65+	70+
Green – Core Support	**1+**	**4** - 9	**10** - 17	**12** - 19	**17** - 32	**25** - 49	**26** - 49	**36** - 62	**43** - 69	**40** - 59	**50** - 64	**55** - 69
Yellow – Strategic Support	0	1 - 3	6 - 9	8 - 11	14 - 16	17 - 24	18 - 25	23 - 35	27 - 42	30 - 39	40 - 49	47 - 54
Red – Intensive Support	--	0	0 - 5	0 - 7	0 - 13	0 - 16	0 - 17	0 - 22	0 - 26	0 - 29	0 - 39	0 - 46

Oral Reading Fluency (ORF) Words Correct

Kindergarten to Third Grade

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Blue – Core Support				35+	57+	76+	85+	117+	128+	105+	141+	136+
Green – Core Support				**10** - 34	**21** - 56	**39** - 75	**49** - 84	**78** - 116	**94** - 127	**73** - 104	**105** - 140	**114** - 135
Yellow – Strategic Support				5 - 9	10 - 20	26 - 38	29 - 48	59 - 77	77 - 93	55 - 72	85 - 104	96 - 113
Red – Intensive Support				0 - 4	0 - 9	0 - 25	0 - 28	0 - 58	0 - 76	0 - 54	0 - 84	0 - 95

Fourth Grade to Eighth Grade

Support	Fourth Beg	Fourth Mid	Fourth End	Fifth Beg	Fifth Mid	Fifth End	Sixth Beg	Sixth Mid	Sixth End	Seventh Beg	Seventh Mid	Seventh End	Eighth Beg	Eighth Mid	Eighth End
Blue Core Support	131+	159+	159+	139+	149+	157+	151+	157+	160+	152+	161+	164+	142+	156+	159+
Green Core Support	**87** - 130	**121** - 158	**125** - 158	**103** - 138	**122** - 148	**137** - 156	**123** - 150	**133** - 156	**141** - 159	**126** - 151	**136** - 160	**141** - 163	**125** - 141	**131** - 155	**135** - 158
Yellow Strategic Support	62 - 86	98 - 120	99 - 124	81 - 102	108 - 121	124 - 136	99 - 122	117 - 132	125 - 140	101 - 125	121 - 135	127 - 140	110 - 124	116 - 130	121 - 134
Red Intensive Support	0 - 61	0 - 97	0 - 98	0 - 80	0 - 107	0 - 123	0 - 98	0 - 116	0 - 124	0 - 100	0 - 120	0 - 126	0 - 109	0 - 115	0 - 120

©2026 University of Oregon. All rights reserved.

Oral Reading Fluency (ORF) Accuracy

First Grade to Third Grade

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Green – Core Support				**67+**	**87+**	**91+**	**92+**	**96+**	**96+**	**96+**	**96+**	**96+**
Yellow – Strategic Support				41 - 66	54 - 86	85 - 90	84 - 91	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95
Red – Intensive Support				0 - 40	0 - 53	0 - 84	0 - 83	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90

Fourth Grade to Eighth Grade

Support	Fourth Beg	Fourth Mid	Fourth End	Fifth Beg	Fifth Mid	Fifth End	Sixth Beg	Sixth Mid	Sixth End	Seventh Beg	Seventh Mid	Seventh End	Eighth Beg	Eighth Mid	Eighth End
Green Core Support	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**	**96+**
Yellow Strategic Support	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95	91 - 95
Red Intensive Support	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90	0 - 90

©2026 University of Oregon. All rights reserved.

Maze

Second Grade to Third Grade

Support	Kinder Beg	Kinder Mid	Kinder End	First Beg	First Mid	First End	Second Beg	Second Mid	Second End	Third Beg	Third Mid	Third End
Blue – Core Support							11.0+	14.5+	18.0+	15.0+	20.5+	22.5+
Green – Core Support							**5.0** - 10.5	**9.0** - 14.0	**9.5** - 17.5	**8.0** - 14.5	**12.0** - 20.0	**15.5** - 22.0
Yellow – Strategic Support							2.5 - 4.5	6.5 - 8.5	7.0 - 9.0	5.0 - 7.5	9.5 - 11.5	12.0 - 15.0
Red – Intensive Support							0 - 2.0	0 - 6.0	0 - 6.5	0 - 4.5	0 - 9.0	0 - 11.5

Fourth Grade to Eighth Grade

Support	Fourth Beg	Fourth Mid	Fourth End	Fifth Beg	Fifth Mid	Fifth End	Sixth Beg	Sixth Mid	Sixth End	Seventh Beg	Seventh Mid	Seventh End	Eighth Beg	Eighth Mid	Eighth End
Blue Core Support	21.0+	23.5+	28.0+	20.0+	27.0+	29.5+	23.0+	30.5+	33.5+	25.5+	33.0+	38.5+	24.5+	32.0+	38.0+
Green Core Support	**14.5** - 20.5	**16.5** - 23.0	**17.0** - 27.5	**13.5** - 19.5	**17.0** - 26.5	**21.0** - 29.0	**14.5** - 22.5	**19.5** - 30.0	**26.5** - 33.0	**20.0** - 25.0	**24.5** - 32.5	**29.5** - 38.0	**20.0** - 24.0	**26.0** - 31.5	**28.0** - 37.5
Yellow Strategic Support	11.0 - 14.0	13.0 - 16.0	14.0 - 16.5	10.5 - 13.0	14.5 - 16.5	18.0 - 20.5	12.5 - 14.0	15.0 - 19.0	20.5 - 26.0	15.5 - 19.5	18.0 - 24.0	24.5 - 29.0	16.5 - 19.5	19.5 - 25.5	24.5 - 27.5
Red Intensive Support	0 - 10.5	0 - 12.5	0 - 13.5	0 - 10.0	0 - 14.0	0 - 17.5	0 - 12.0	0 - 14.5	0 - 20.0	0 - 15.0	0 - 17.5	0 - 24.0	0 - 16.0	0 - 19.0	0 - 24.0

©2026 University of Oregon. All rights reserved.

Appendix F: Composite Score Derivation and Calculation Guide

Derivation of the DIBELS 8 Composite Score

The DIBELS 8 Composite score is a linear combination of scores on DIBELS 8 measures that provides an estimate of overall student literacy skill. To compute composite scores for DIBELS 8th Edition, we used a Confirmatory Factor Analysis (CFA) approach. For each grade, we tested a series of theory-based, one-factor reading models based on theories on literacy development and literacy assessment. The models were built iteratively, starting with a base model for each grade, where all DIBELS 8 measures for that grade were loaded on the common reading factor. See Table F.1 for a summary of measures by grade. Then, this model was extended by modeling different types of covariances. Table F.2 on page 137 presents the theoretical reading factor models that were tested by grade level.

Table F.1. DIBELS 8 Measures Available by Grade

Grade	LNF	PSF	NWF	WRF	ORF	Maze
K	X	X	X	X		
1	X	X	X	X	X	
2			X	X	X	X
3			X	X	X	X
4					X	X
5					X	X
6					X	X
7					X	X
8					X	X

©2026 University of Oregon. All rights reserved.

Table F.2. Theoretical Reading Factor Models by Grade

Grade	Model
K-3	All available DIBELS 8 scores
	All available DIBELS 8 scores + ORF-WRC – ORF-ACC covariance
	All available DIBELS 8 scores + NWF-CLS – NWF-WRC covariance
	All available DIBELS 8 scores + ORF-WRC – WRF covariance
	All available DIBELS 8 scores + WRF – NWF-WRC covariance
	All available DIBELS 8 scores + ORF-WRC – Maze covariance
4-8	All available DIBELS 8 scores (ORF-WRC, ORF-Acc, Maze)
	All available DIBELS 8 scores + all covariances

In the reading factor models for grades K–3, the models that include covariances between ORF-WRC and ORF-ACC and between NWF-CLS and NWF-WRC take into account the residuals that arise from including multiple scores from the same subtest in the model. Modeling the covariance between ORF-WRF and WRF takes into account residuals associated with including multiple subtests that measure real word reading, while modeling the covariance between WRF and NWF-WRC accounts for the residuals associated with including multiple subtests that measure blending of sounds into words. Finally, modeling the covariance between ORF-WRC and Maze takes into account the residuals associated with including multiple subtests that measure reading comprehension.

The final model for each grade level was determined by comparing model fits. Model fit was evaluated using the CFI (Bentler, 1990; acceptable fit ≥.95), root mean square error of approximation (RMSEA; Browne & Cudeck, 1993; acceptable fit ≤.06), the standardized root mean square residual (RMSR; Hu & Bentler, 1999; acceptable fit ≤.10), Akaike information criterion (AIC; Burnham & Anderson, 2004; lower values, relative to other nested models, are better), and Bayesian information criterion (BIC; Burnham & Anderson, 2004; lower values, relative to other nested models, are better). Models were fit to data collected in the fall of 2018, using maximum likelihood estimation.

In grades K-3, the resulting best-fitting model included all available DIBELS 8 measures for each grade level and the covariance between NWF-CLS and NWF-WRC. In grades 4-8, the best-fitting model included all the available DIBELS 8 measures but no covariances. Unstandardized factor loadings (i.e.,

©2026 University of Oregon. All rights reserved.

weights) in the final reading models were all statistically significant.

To verify that the weights derived from analyses using data from fall 2018 generalized to other, previous samples, we also conducted a series of parallel analyses using data from fall, winter, and spring of 2017-18. The weights derived for each season using the 2017-18 data were very similar to both each other, and to the weights derived using the fall 2018 data, giving us confidence in the generalizability of the weights estimated using the fall 2018 data to other seasons and other samples.

Calculating the DIBELS 8 Composite Score

To calculate the DIBELS 8 composite score, a student must have been administered all available subtests for their grade. Apply the following steps, in order:

1. For each subtest raw score, multiply the student's raw score by the Weight listed in the table on the next page, rounding the result to the 100^{ths} place.

If a student does not have a subtest raw score due to the Discontinue or Gating Rules, use the constant from the table in the next section on page 142 for the missing subtest scores.

2. Sum the resulting weighted scores across all applicable subtests.
3. From that sum, subtract the Mean for the appropriate grade from the table on the next page.
4. Divide the result by the standard deviation (SD) for the appropriate grade in the table on the next page.
5. Multiply the result by 40.
6. Add the scaling Constant corresponding to the grade and season in which the student was tested from the table on the next page and round to the ones place. The result is the composite score.

Note that ORF Accuracy should be represented in these calculations as a proportion of words correct (e.g., 0.99), rather than percent correct (e.g., 99%).

The weight applied at each grade for each measure is consistent across all time periods with the exception of kindergarten. In kindergarten at Beginning of Year (BOY) the weight of LNF is greater

©2026 University of Oregon. All rights reserved.

than it is at the Middle of Year (MOY) and End of Year (EOY).

You can find DIBELS 8th Edition Composite Score Calculation Assistance spreadsheet and the Technical Report on Dimentionality of the DIBELS 8 Composites Score on the Research page on the DIBELS website: https://dibels.uoregon.edu/research.

Example calculations are provided at the end of this Appendix.

Grade	Subtest score	Weight	Mean	SD	Fall constant	Winter constant	Spring constant
Kindergarten	LNF BOY	35.44	729	630	289	364	398
	LNF MOY/EOY	8.86					
	PSF	4.13					
	NWF-CLS	14.93					
	NWF-WRC	3.56					
	WRF	5.62					
First	LNF	10.72	3371	2251	360	400	440
	PSF	2.13					
	NWF-CLS	23.13					
	NWF-WRC	7.79					
	WRF	13.51					
	ORF-WRC	25.36					
	ORF-ACC	0.25					
Second	NWF-CLS	32.74	7085	3811	360	400	440
	NWF-WRC	10.95					
	WRF	21.26					
	ORF-WRC	35.36					
	ORF-ACC	0.15					
	MAZE	4.28					

©2026 University of Oregon. All rights reserved.

Grade	Subtest score	Weight	Mean	SD	Fall constant	Winter constant	Spring constant
Third	NWF-CLS	40.02	10051	4349	360	400	440
	NWF-WRC	11.80					
	WRF	19.83					
	ORF-WRC	39.42					
	ORF-ACC	0.09					
	MAZE	4.79					
Fourth	ORF-WRC	36.42	4563	1771	360	400	440
	ORF-ACC	0.06					
	MAZE	6.29					
Fifth	ORF-WRC	31.12	4085	1299	360	400	440
	ORF-ACC	0.03					
	MAZE	4.58					
Sixth	ORF-WRC	40.71	6087	1685	360	400	440
	ORF-ACC	0.05					
	MAZE	5.03					
Seventh	ORF-WRC	40.55	6444	1960	360	400	440
	ORF-ACC	0.06					
	MAZE	7.34					
Eighth	ORF-WRC	37.69	4824	1506	360	400	440
	ORF-ACC	0.03					
	MAZE	6.75					

Composite Score Using Discontinue and Gating Rules

Special rules apply to calculate the DIBELS 8 composite score when a discontinue or gating rule has been used and a student has not been administered all available subtests for their grade. In these cases a constant is used in place of the missing raw scores. The steps are the same as when

©2026 University of Oregon. All rights reserved.

calculating a regular composite score with the additional step of substituting a constant for the missing values. Example calculations are provided after the tables on the following pages.

Discontinue Rules

There are three discontinue rules that impact calculation of a composite score. In all cases, when a measure is discontinued a score of zero is given for that measure and the remaining measures are not administered.

Grade K. Beginning of Year (BOY) - If PSF is discontinued, do not administer NWF and WRF. Enter zero for PSF. Do not enter scores for the remaining subtests: NWF and WRF.

Grade K. Middle of Year (MOY) - If NWF is discontinued, do not administer WRF. Enter zero for NWF. Do not enter scores for the remaining subtest: WRF.

Grade 1. Beginning of Year (BOY) - If WRF is discontinued, do not administer ORF. Enter zero for WRF. Do not enter scores for the remaining subtest: ORF

Gating Rules

There are eight time periods across three grades where a gating rule impacts calculation of a composite score. In these cases, when a measure is not given because a student's scores are in the blue negligible risk category on the specified measures, a constant value can be used to calculate a composite score.

Grade 1. Middle of Year (MOY) and End of Year (EOY) - If the student scores in the blue range on NWF-CLS, then you do not have to administer LNF or PSF. Leave those scores blank if not administered.

Grade 2 and 3. Beginning of Year (BOY), Middle of Year (MOY), and End of Year (EOY) - If the student scores in the blue range on ORF-WRC, then you do not have to administer NWF or WRF. Leave those scores blank if not administered. Always give Maze.

©2026 University of Oregon. All rights reserved.

Composite Score Constants for Discontinued and Gated Administrations

	Grade	Time Period	LNF	PSF	NWF CLS	NWF WRC	WRF	ORF WRC	ORF ACC
Discontinue	Kindergarten	BOY (fall)			0	0	0		
		MOY (winter)					0		
	First	BOY (fall)						0	0
Gating	First	MOY (winter)	66	56					
		EOY (spring)	68	60					
	Second	BOY (fall)			85	24	49		
		MOY (winter)			102	35	62		
		EOY (spring)			116	38	69		
	Third	BOY (fall)			120	33	59		
		MOY (winter)			137	45	64		
		EOY (spring)			140	44	69		

©2026 University of Oregon. All rights reserved.

Example with Full Data

For a second-grade student with fall DIBELS 8 scores of 152 for NWF Correct Letter Sounds (CLS), 48 for NWF Words Read Correctly (WRC), 45 for WRF, 88 for ORF Words Read Correctly (WRC), 99% ORF Accuracy, and 11.0 for Maze Adjusted, we would calculate this student's composite score as follows.

Step 1: Multiply each subtest raw score by the corresponding weight listed in the table. Round Weighted scores to the 100ths place.

Subtest score	Raw score	Weight	Weight score
NWF-CLS	152	* 32.74	= 4976.48
NWF-WRC	48	* 10.95	= 525.60
WRF	45	* 21.26	= 956.70
ORF-WRC	88	* 35.36	= 3111.68
ORF-ACC	0.99	* 0.15	= 0.15
Maze	11	* 4.28	= 47.08

Step 2: Sum the resulting weighted scores across all applicable subtests:

4976.48 + 525.6 + 956.7 + 3111.68 + 0.15 + 47.08 = 9617.69

Step 3: Subtract from that sum the mean of the weighted scores for the appropriate grade:

9617.69 – 7085 = 2532.69

Step 4: Divide that value by the standard deviation for the appropriate grade:

2532.69 / 3811 = 0.6646

Step 5: Multiply that score by 40:

0.66 * 40 = 26.5829

Step 6: Add the scaling constant corresponding to the season in which the student was tested and round to the ones place to obtain the final composite score:

26.5829 + 360 = 386

©2026 University of Oregon. All rights reserved.

Example with Discontinue Rule Implemented in Kindergarten

For a kindergarten student with Beginning of Year (BOY/fall) DIBELS 8 scores of 10 for LNF, 0 for PSF, and no scores for NWF-CLS, NWF-WRC, or WRF due to the discontinue rule, we would calculate this student's composite score as follows.

Step 1: Multiply each subtest raw score by the corresponding weight listed in the table. Use a zero for the missing subtest/s score/s.

Subtest score	Raw score	Weight	Weight score
LNF	10	* 35.44	= 354.40
PSF	0	* 4.13	= 0.00
NWF-CLS	0	* 14.93	= 0.00
NWF-WRC	0	* 3.56	= 0.00
WRF	0	* 5.62	= 0.00

Step 2: Sum the resulting weighted scores across all applicable subtests:

354.40 + 0.00 + 0.00 + 0.00 + 0.00 = 354.40

Step 3: Subtract from that sum the mean of the weighted scores for the appropriate grade:

354.40 – 729 = -374.60

Step 4: Divide that value by the standard deviation for the appropriate grade:

-374.60 / 630 = -0.5946

Step 5: Multiply that score by 40:

-0.5946 * 40 = -23.7841

Step 6: Add the scaling constant corresponding to the season in which the student was tested and round to the ones place to obtain the final composite score:

-23.7841 + 289 = 265

©2026 University of Oregon. All rights reserved.

Example with Gating Rule Implemented in Second Grade

For a second grade student with Beginning of Year (BOY/fall) DIBELS 8 scores of 93 for ORF-WRC, 0.99 (99%) ORF-ACC, 11.5 for Maze Adjusted, and no scores for NWF-CLS, NWF-WRC, or WRF due to the gating rule, we would calculate this student's composite score as follows.

Step 1: Multiply each subtest raw score by the corresponding weight listed in the table. Use the stant values for the missing subtest/s score/s.

Subtest score	Raw score	Weight	Weight score
NWF-CLS	85	* 32.74	= 2782.90
NWF-WRC	24	* 10.95	= 262.80
WRF	49	* 21.26	= 1041.74
ORF-WRC	93	* 35.36	= 3288.48
ORF-ACC	0.99	* 0.15	= 0.15
Maze	11.5	* 4.28	= 49.22

Step 2: Sum the resulting weighted scores across all applicable subtests:

2782.90 + 262.80 + 1041.74 + 3288.48 + 0.15 + 49.22 = 7425.29

Step 3: Subtract from that sum the mean of the weighted scores for the appropriate grade:

7425.29 – 7085 = 340.29

Step 4: Divide that value by the standard deviation for the appropriate grade:

340.29 / 3811 = 0.0893

Step 5: Multiply that score by 40:

0.0893 * 40 = 3.5716

Step 6: Add the scaling constant corresponding to the season in which the student was tested and round to the ones place to obtain the final composite score:

3.5716 + 360 = 364

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Kindergarten

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
LNF	_____	x 35.44 if Beginning of year x 8.86 if Middle or End of year	= _____
PSF	_____	x 4.13	= _____
NWF-CLS	_____	x 14.93	= _____
NWF-WRC	_____	x 3.56	= _____
WRF	_____	x 5.62	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ **- 729 =** _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ **÷ 630 =** _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ **x 40 =** _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 289, Winter/Middle = 364, Spring/End = 398.

_____ **+** _____ **=** _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet First Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
LNF	_____	x 10.72	= _____
PSF	_____	x 2.13	= _____
NWF-CLS	_____	x 23.13	= _____
NWF-WRC	_____	x 7.79	= _____
WRF	_____	x 13.51	= _____
ORF-WRC	_____	x 25.36	= _____
ORF-ACC	_____	x 0.25	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ - 3371 = _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ ÷ 2251 = _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ x 40 = _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ + _____ = _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Second Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
NWF-CLS	_____	x 32.74	= _____
NWF-WRC	_____	x 10.95	= _____
WRF	_____	x 21.26	= _____
ORF-WRC	_____	x 35.36	= _____
ORF-ACC	_____	x 0.15	= _____
Maze	_____	x 4.28	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ - 7085 = _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ ÷ 3811 = _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ x 40 = _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ + _____ = _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Third Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
NWF-CLS	_____	x 40.02	= _____
NWF-WRC	_____	x 11.80	= _____
WRF	_____	x 19.83	= _____
ORF-WRC	_____	x 39.42	= _____
ORF-ACC	_____	x 0.09	= _____
Maze	_____	x 4.79	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ - 10051 = _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ ÷ 4349 = _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ x 40 = _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ + _____ = _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Fourth Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
ORF-WRC	_____	x 36.42	= _____
ORF-ACC	_____	x 0.06	= _____
Maze	_____	x 6.29	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ - 4563 = _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ ÷ 1771 = _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ x 40 = _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ + _____ = _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Fifth Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
ORF-WRC	_____	x 31.12	= _____
ORF-ACC	_____	x 0.03	= _____
Maze	_____	x 4.58	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ **- 4085 =** _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ **÷ 1299 =** _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ **x 40 =** _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ **+** _____ **=** _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Sixth Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
ORF-WRC	_____	x 40.71	= _____
ORF-ACC	_____	x 0.05	= _____
Maze	_____	x 5.03	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ - 6087 = _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ ÷ 1685 = _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ x 40 = _____(

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ + _____ = _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Seventh Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
ORF-WRC	_____	x 40.55	= _____
ORF-ACC	_____	x 0.06	= _____
Maze	_____	x 7.34	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ - 6444 = _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ ÷ 1960 = _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ x 40 = _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ + _____ = _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

DIBELS 8th Edition Composite Score Calculation Worksheet Eighth Grade

Step 1. Multiply each subtest raw score by the weight listed. Round Weighted scores to the 100ths place.

Subtest	Raw score	Weight	Weighted score
ORF-WRC	_____	x 37.69	= _____
ORF-ACC	_____	x 0.03	= _____
Maze	_____	x 6.75	= _____

Step 2. Sum the weighted scores from Step 1. Total = _____

Step 3. Subtract the mean of the weighted score from the sum of the weighted scores.

_____ - 4824 = _____

(Total from Step 2)

Step 4. Divide value from Step 3 by standard deviation.

_____ ÷ 1506 = _____

(Value from Step 3)

Step 5. Multiply value from Step 4 by 40.

_____ x 40 = _____

(Value from Step 4)

Step 6. Add the scaling constant for the season in which the student was tested and round to the ones place to obtain the final composite score.

Constants: Fall/Beginning = 360, Winter/Middle = 400, Spring/End = 440.

_____ + _____ = _____

(Value from Step 5) (constant) (final composite score)

©2026 University of Oregon. All rights reserved.

Index

A

Abbreviations —79

Acceptable prompts —53, 57, 66, 74, 77

Accommodations —48, 49

Accurate —45

Additions —59

Administration —9, 38, 39, 43, 49, 52, 55, 56, 63, 64, 72, 73, 75, 76, 80, 81, 87, 126

Appropriate Uses —12, 23

Assistive technology —49

B

Benchmark —36, 41, 42, 52, 53, 56, 58, 64, 66, 73, 74, 76, 77, 81, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 133

C

Calculating —141

Calculation —139, 149, 150, 151, 152, 153, 154, 155, 156, 157

Cautions —93

Checklist —126, 127, 128, 129, 130, 131

Choosing —44, 95

Classification —90

comparisons —38, 92

Composite —89, 133, 134, 139, 141, 144, 145, 149, 150, 151, 152, 153, 154, 155, 156, 157

Composite Scores —89

Considerations —45

Core Support —133, 134, 135, 136, 137, 138

Correct responses —54, 58, 67, 74, 77, 84

©2026 University of Oregon. All rights reserved.

cut-score —22, 42, 90, 91, 98

D

data system —50, 92

Derivation —139

Development —7, 22

DIBELS —6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 32, 33, 35, 36, 37, 38, 39, 40, 42, 43, 45, 46, 47, 48, 49, 50, 84, 89, 90, 91, 92, 93, 94, 95, 97, 98, 125, 131, 133, 139, 140, 141, 142, 144, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157

Discontinue —14, 40, 41, 53, 57, 58, 66, 74, 77, 83, 141, 144, 145, 147

Discontinue Rule —41, 147

Dyslexia —8, 14, 92

E

Eighth Grade —157

evaluating the progress —92

Example —19, 125, 142, 144, 146, 147, 148

F

Fidelity —55, 63, 72, 75, 80, 87, 126, 127, 128, 129, 130, 131

Fifth Grade —154

First Grade —41, 150

Font —14, 15, 16, 32, 38, 54, 102, 104, 105, 108

Fourth Grade —153

Frequency —96, 97

G

Gating —40, 42, 141, 144, 145, 148

Gating Rule —42, 148

©2026 University of Oregon. All rights reserved.

Goals —133

Guide —48, 125, 139

Guidelines —38, 39

H

History —6

I

Inconsistent marking —85

Incorrect responses —54, 58, 67, 74, 78, 83, 84

Inserted Sounds —71

Insertions —78

Instructions —38, 45

Intensive —133, 134, 135, 136, 137, 138

Intensive Support —133, 134, 135, 136, 137, 138

Interpreting —89, 93

K

Kindergarten —16, 41, 142, 145, 147, 149

L

Letter Naming Fluency —6, 7, 9, 17, 24, 52, 126, 134

Letter reversals —54

Letter sounds —55, 68, 70

LNF —7, 9, 52, 92, 134

M

Materials —52, 56, 64, 73, 76, 81

Maze —7, 8, 9, 12, 15, 17, 21, 22, 35, 36, 38, 39, 43, 44, 49, 81, 83, 84, 87, 91, 92, 96, 97, 118, 119, 120, 121, 122, 123, 124, 131, 138, 139, 140, 145, 146, 148, 151, 152,

©2026 University of Oregon. All rights reserved.

153, 154, 155, 156, 157

Mispronounced segment —61

Mispronounced words —80

Monitoring —95, 96, 97, 98, 103, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124

Multiple marks with clear intention —87

N

Nonsense Word Fluency —7, 10, 18, 26, 48, 64, 128, 135

No segmentation —62

NWF —7, 8, 10, 11, 14, 16, 18, 19, 23, 26, 27, 39, 40, 41, 42, 45, 47, 49, 58, 64, 66, 67, 72, 89, 92, 93, 95, 97, 135, 139, 140, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152

O

Omissions —55, 63, 69, 72, 75, 80

Oral Reading Fluency —7, 11, 15, 20, 31, 32, 76, 130, 136, 137

ORF —7, 8, 11, 12, 14, 15, 16, 17, 20, 21, 23, 31, 33, 35, 39, 40, 41, 42, 44, 45, 49, 74, 76, 77, 80, 89, 91, 92, 95, 96, 97, 98, 110, 111, 112, 113, 114, 115, 116, 117, 136, 137, 139, 140, 141, 142, 143, 144, 145, 146, 148, 150, 151, 152, 153, 154, 155, 156, 157

Overlapping segmentation —61

P

Partially correct responses —69

Partial segmentation —60

Passage —32, 36, 78, 79, 80, 81, 101, 110, 118

Percentile Ranks —91

Phonemic Segmentation Fluency —6, 7, 10, 18, 25, 56, 127

©2026 University of Oregon. All rights reserved.

Procedures —38

Progress Monitoring —95, 96, 97, 98, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124

prompts —38, 50, 53, 57, 66, 74, 77, 83, 126, 127, 128, 129, 130, 131

Pronunciation —48, 125

PSF —7, 8, 10, 14, 18, 21, 23, 25, 26, 39, 41, 42, 45, 49, 56, 57, 58, 63, 92, 93, 95, 97, 135, 139, 142, 144, 145, 147, 149, 150

R

Ranks —91

Raw scores —91, 92

R-controlled vowels —47

R-Controlled vowels —47, 62

Reading —7, 11, 15, 20, 28, 31, 32, 73, 76, 83, 118, 129, 130, 136, 137, 140

reliability —20, 33, 94, 98

Repeated sounds —68

Repetitions —78

Response to Intervention —97

Risk —90, 97, 133

Rule —41, 42, 53, 57, 58, 66, 74, 77, 147, 148

S

Schwa sounds —59

Scores —49, 50, 89, 93

Scoring —38, 39, 41, 42, 45, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 83, 126

Screening —92

Second Grade —148, 151

Self-corrections —54, 58, 67, 75, 78, 84

©2026 University of Oregon. All rights reserved.

Seventh Grade —156

Sixth Grade —155

Skipped items —12, 86

Sounded out words —75, 79

Sound elongation —60

Sounds followed by correct word —67

Sounds followed by incorrect word —68

Sounds out of order —70

Spelling —19, 62

Statistics —110, 118

Strategic —31, 133, 134, 135, 136, 137, 138

Strategic Support —133, 134, 135, 136, 137, 138

Subtest —9, 16, 95, 142, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157

Support —3, 133, 134, 135, 136, 137, 138

T

Third Grade —19, 152

U

Unclear or multiple responses —87

W

Word order —75, 80

Word Reading Fluency —7, 11, 20, 28, 73, 129, 136

Worksheet —149, 150, 151, 152, 153, 154, 155, 156, 157

WRF —7, 8, 11, 14, 17, 20, 23, 28, 39, 41, 42, 45, 49, 58, 66, 73, 74, 75, 95, 97, 136, 139, 140, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152

Z

Zones of Growth —89, 92, 163

©2026 University of Oregon. All rights reserved.