

Assessing the Readability of *DIBELS® AD Oral Reading Fluency and Daze*

Kelly A. Powell-Smith

Roland H. Good, III

Elizabeth N. Dewey

Rachael J. Latimer

Dynamic Measurement Group

Technical Report #16

Recommended Citation:

Powell-Smith, K., Good, R. H., III, & Dewey, E. N., & Latimer, R. J. (2014). *Assessing the Readability of DIBELS® AD Oral Reading Fluency and Daze*. (Technical Report No. 16). Eugene, OR: Dynamic Measurement Group.

Author Note

The authors thank the faculty, staff, students, and parents of the participating schools for their effort and support during the course of this study. Correspondence regarding this manuscript should be addressed to Kelly A. Powell-Smith, Dynamic Measurement Group, 859 Willamette Street, Suite 320, Eugene, OR 97401; Email: kpowellsmith@dibels.org.

Introduction

DIBELS® are a set of brief measures of early reading that have been found to be reliable and valid for assessing the acquisition of early literacy skills of children who are learning to read in English in the United States. The *DIBELS* measures map on to the critical early reading skills identified by the National Reading Panel (2002) and include indicators of phonemic awareness, alphabetic principle, vocabulary and oral language development, accuracy and fluency with connected text, and comprehension. *DIBELS* data are collected routinely for many schools as part of ongoing, school improvement efforts in reading. To maximize these efforts, refinement of the *DIBELS* measures has occurred. This refinement included revisions of the oral reading fluency (ORF) passages. Several studies have addressed readability of oral reading fluency passages (e.g., Ardoin, Suldo, Witt, Aldrich, & McDonald, 2005; Compton, Appleton, & Hosp 2004; Hintze & Christ, 2004; Powell-Smith & Bradley-Klug, 2001). More recently studies have attempted to equate ORF passages (e.g., Betts, Pickart, & Heistad, 2009; Francis, Santi, Barr, Fletcher, Varisco, Foorman, 2008) due to concerns related to passage variability. Our previous work on *DIBELS Next* used a multi-method approach to understanding passage variability and assigning reading passages for benchmark and progress monitoring purposes (see Powell-Smith, Good, & Atkins, 2010). This study takes a similar approach to examining newly-created *DIBELS AD* ORF and Daze passages.

Purpose of the Study

The purpose of this study was to examine the readability of newly written *DIBELS* Oral Reading Fluency (DORF) and Daze passages for a secure version of spring/end-of-year benchmark assessment called *DIBELS for Accountability Decisions (DIBELS AD)*. *DIBELS AD* has

been developed for use by schools that 1) rely on *DIBELS Next*, 2) are now required to provide teacher evaluation information based on their formative reading assessment data, and 3) wish to have a measure with the reliability, validity, and utility of DIBELS. *DIBELS AD* involves the use of a secure, alternate form of the *DIBELS Next* measures for the end-of-year (spring) benchmark testing. *DIBELS AD* is an add-on component for *DIBELS Next* that can be used as one indicator of student growth within a comprehensive teacher evaluation and accountability program. *DIBELS AD* results can be treated as *DIBELS Next* results for the purpose of making instructional decisions, but can also be used for making accountability decisions. With these purposes in mind, our goal was to identify reading passages comparable to *DIBELS Next* passages that could be used for new secure forms each year of *DIBELS AD*. Thus, in this study any passages that were not aligned to the *DIBELS Next* spring benchmark passages at the same grade level (in terms of average DORF and Daze scores) were revised so that they are comparable to the *DIBELS Next* spring benchmark passages at the same grade level or were replaced with other passages meeting our alignment criteria.

In this technical report we detail the findings from this descriptive study. The specific research questions examined in the study included:

1. What are the mean DORF & Daze passages at each grade level?
2. What are the most extreme passages at each grade level relative to the mean passage(s)?
3. How do the new passage scores compare to scores on current *DIBELS Next* spring benchmark materials?

4. Are there common patterns in the passage characteristics/qualitative features of the outlier passages (e.g., expository vs. narrative text, unusual proper names, sentences with multiple clauses, etc.)?
5. What is the alternate-form reliability of the new DORF and Daze passages?
6. How should DORF passages be formed into triads and then paired with a Daze passage?

In this report, we provide information about the passages under review relative to these research questions. Both descriptive and correlational data are presented. Additionally, this report describes the process by which the passages were placed in groups of three (triads) for end-of-year benchmark assessment. For further information regarding the development of these measures, as well as additional information about *DIBELS AD*, please visit:

<https://dibels.org/>.

Method

Participants

The participants in this study were 268 students in first through sixth grade in three elementary schools from three school districts in three states in the Midwest Region of the United States (United States Census Bureau). During the study, data were collected at schools, and student test booklets were boxed and shipped to DMG offices. However, data from one of the three schools did not arrive in our offices before the analysis deadline, and therefore could not be included in the analysis. The initial sample size examined included 179 student booklets.

Demographic data on both schools included in the data analysis is found in Table 1. One of the schools was a private school and one school was a public school. One school was located in a remote rural area and one in a large suburban area. The schools ranged in the size of the

student population served from 392 to 731 students and the grade levels served ranged from PK to grade 8. Information on the percentage of students participating in the federal free/reduced price lunch program and Title 1 eligibility was available for only one of the two schools. The student population in both of the schools was primarily White.

The percent of students in our sample who were At/Above Benchmark on *DIBELS Next* DORF and Daze compared to winter and spring benchmark expectations at each grade level is reported in Table 2. These data show that most, and in some cases all, students earned scores at or above the benchmark goal for winter. Between 60% and 80% of students earned scores that were at or above the spring benchmark goal. It should be noted that testing occurred in February, March, and April. In general, this timeframe falls between when winter (middle-of-year) and spring (end-of-year) benchmark data collection would occur.

Measures

Daze and DORF reading passages for this study were created by a team of authors following specific reading passage development criteria and story ideas. All passages were leveled by grade using the same procedures used in the development of the *DIBELS Next* DORF passages. Information on the readability and technical characteristics (based upon DMG's Passage Difficulty Index) of the *DIBELS Next* passages is detailed in Powell-Smith, Good, & Atkins (2010). Characteristics of the *DIBELS AD* DORF and Daze passages based upon the DMG Passage Difficulty Index are found in the Appendix (Tables A1 - A10).

***DIBELS* Oral Reading Fluency (DORF).** *DIBELS* Oral Reading Fluency (DORF) is a measure of advanced phonics and word attack skills, accurate and fluent reading of connected text, and reading comprehension. The DORF passages and procedures are based on the program of

research and development of Curriculum-Based Measurement (CBM) of reading by Stan Deno and colleagues at the University of Minnesota (Deno, 1985; Shinn, 1989). There are two components to DORF. The first part is oral reading fluency and the second part is passage retell. For the purposes of this study, Retell was not administered. For the oral reading fluency component, students are given an unfamiliar, grade-level text passage and asked to read aloud for 1 minute. The assessor listens and follows along using an assessor copy of the passage. The assessor marks the errors made by the student including: substitutions, omissions, and hesitations for more than 3 seconds. DORF is scored by determining the number of words read correctly and the number of errors. The student's accuracy rate is calculated based on the number of words read correctly and the number of errors, using the following formula:

$$Accuracy = 100 \times \frac{\text{median words correct}}{\text{median words correct} + \text{median errors}}$$

A series of studies examined the technical adequacy of CBM ORF procedures in general. Test-retest reliabilities for elementary-aged students ranged from .92 to .97; alternate-form reliability of different reading passages drawn from the same level ranged from .89 to .94 (Tindal, Marston, & Deno, 1983). Criterion-related validity data from eight separate studies in the 1980s reported coefficients ranging from .52 to .91 (Good & Jefferson, 1998).

DIBELS-Maze (Daze). Daze is the standardized, *DIBELS* version of a maze testing procedure for measuring reading comprehension. The purpose of a maze assessment is to measure the reasoning processes that underlie comprehension. Specifically, Daze assesses the student's ability to construct meaning from text, word recognition skills, background information and

prior knowledge, familiarity with linguistic properties such as syntax and morphology, and reasoning skills. Daze can be administered to a whole class at the same time, to a small group of students, or individually. Students are given a passage in which approximately every seventh word has been replaced by a box containing the correct word and two distractor words. Using standardized directions, students are asked to read the passage silently and circle their word choices. The student receives credit for selecting the word that best fits the passage from the choices available in each box presented in the reading passage. The scores that are recorded are the number of correct and incorrect responses. An adjusted score, which compensates for guessing, is calculated based on the number of correct and incorrect responses. The following procedure is used to calculate the adjusted score:

$$\text{Daze Adjusted Score} = \text{number of correct responses} - (\text{number of incorrect responses} \div 2).$$

The result of this calculation is rounded to the nearest whole number, with half-points (0.5) rounded up. If the adjusted score is negative, a score of 0 is recorded. Since there are three possible responses for each item, the adjusted score expected for a student who guesses on every item is 0. Scores are not pro-rated. The maximum score a student can receive is equal to the number of items on the form.

Multiple studies have been conducted evaluating the reliability and validity of Daze. Here, we report the median alternate-form reliability and validity where multiple coefficients were available. This information is also available in the *DIBELS Next Technical Manual*. Daze adjusted score alternate-form reliability for grades 3 through 6 ranges from .66 to .81, and inter-rater reliability ranges from .98 to .99 (Good, Kaminski, Dewey, Wallin, Powell-Smith, and Latimer, 2013). Predictive and concurrent validity with GRADE Total Test (Williams, 2001) range from .58

to .68. The information summarized here and in the *DIBELS Next Technical Manual* suggests that Daze is a good indicator of reading comprehension (see Good et al., 2013 or details).

Procedures

Participant recruitment and selection. Research sites were recruited for this study via email, telephone, or in person. Prior to student recruitment and selection for this study, IRB approval was secured followed by school and/or district approval to conduct the study. General education classroom teachers were recruited with the assistance of the local project coordinator. Students were selected from these teachers' classrooms provided they met the inclusion criteria. Parental consent and student assent was sought prior to any testing being conducted. The school sites, local coordinators, and student participants received compensation for their participation in this study.

Students were eligible to participate if they were receiving English language reading instruction, including students with disabilities and those who are English language learners, provided they had the response capabilities to participate. Each school site selected the *n*th student (e.g., every 5th student) from lists of all students at each grade level meeting the inclusion criteria until 15 students per grade were selected.

Training. Data collectors participated in a one-hour training conducted by DMG research scientists (via webinar) or in person by a *DIBELS* professional development specialist or *DIBELS Next* Mentor. A review of the materials and administration and scoring rules was provided during this time. Examiners were provided with opportunities for practice and to ask questions during the training. During the training, examiners were given sample items from each measure, except Daze, to practice correct scoring procedures. Agreement on scoring and

fidelity to standardize administration and scoring procedures was checked. Prior to data collection, all data collectors had their accuracy of administration and scoring checked by the local site coordinator using the *DIBELS Next* accuracy in implementation checklists. Scores were required to be within 2 points or re-training and re-checking of accuracy was conducted. Data collectors were offered compensation for their time.

Data Collection. Data collection occurred over approximately two to four weeks during February and March of 2013 in the two schools whose data were included in the analyses. All *DIBELS Next* testing was conducted by on-site coordinators, school teachers, or other school personnel who were trained to administer and score the DORF and Daze *DIBELS* measures.

Testing was conducted over approximately 4 to 7 testing sessions. Over the course of these testing sessions, students were administered 16 DORF passages (four passages per session) at their grade level. DORF testing was conducted with students individually. Each testing session was approximately 8 to 10 minutes in length. A discontinue rule was used during the data collection process to prevent student frustration. If a student read fewer than 10 words correct on the first passage administered, they were discontinued from the study. If more than 2 students in grade 1 or 2, or more than 5 in grades 3 - 6, met the discontinue criterion, then another student at that grade level was selected to participate from the pool of eligible students so that the sample would not drop below 40-45 students per grade level. Students in grades 3 – 6 also were administered six Daze passages (two passages in each of three additional sessions). The Daze assessment was group-administered. Each of the sessions was approximately 7 - 10 minutes in length. All passages, both DORF and Daze, were administered in a random order specific for each participating student.

Measure scoring and reliability. The measures were scored by the data collectors on site. All student reading packets were returned to DMG. All data collected was re-scored by DMG personnel before being entered into a spreadsheet for data analysis. Scoring accuracy was checked on 100% of the passages. Any discrepancies in scoring were checked with the principal investigator. Any math errors (e.g., when calculating the number of words read correctly on DORF) were corrected in the student packets and in the data entry files.

Data entry and reliability. Data was entered twice, by separate personnel, into two Excel spreadsheets and then checked for accuracy. Data-entry reliability was checked electronically by comparing the two Excel spreadsheets using an automated comparison formula. Any disagreements in the files resulted in re-examining the student packet to determine the accurate score and reconcile any disagreements in the electronic files. Any data-entry errors (e.g., typos such as entering the same order number for two different passages) were corrected before analysis.

Data management. Detailed data management procedures were followed to assure that a corrected and clean data set was used for data analysis. Data were checked for complete records, invalid scores, and duplicate student ID numbers. No duplicate student IDs were found. In addition to invalid-score checking, passages were chosen randomly within each student booklet to check score accuracy. Any scoring errors were corrected. A few scoring errors were also discovered during data analysis and were corrected. Data were then evaluated for unusual scores that may not represent the intended nature of the task (e.g., if a student guessed on Daze or engaged the DORF task without reading for meaning thus resulting in a high error rate). The final sample had 159 students with DORF passage data, and 103 students with Daze

passage data. See measure-specific sections that follow the section on scoring decision rules and outliers for details regarding the way in which these data management steps impacted our final sample sizes by measure and grade.

Scoring decision rules and outliers. We examined all of the data by grade for the presence of outlier students, patterns of guessing, missing scores, and/or invalid scores. We defined outlier students as those students who earned scores at the extreme high end or extreme low end of the scoring range compared with the rest of the students in the sample. The scores on *DIBELS AD* passages were compared directly to the scores on end-of-year *DIBELS Next* passages. Thus, students that were missing any or all scores from their *DIBELS Next* passages were omitted from analysis. What follows is a detailed description of the decision rules for deleting outlier students and/or individual passage scores from each grade for DORF and Daze. Removal of scores for individual passages decreased the sample size for the analyses including that passage, but did not affect the overall sample size of the study.

DIBELS Oral Reading Fluency. Two scores are recorded for each student's DORF assessment: the number of words read correctly and the number of errors, which indicates the number of words read incorrectly. The sum of these two scores should not be greater than the number of words in the passage. In our sample, there were no students with combined recorded scores that were above the maximum possible score on any individual passage.

Scores for students performing at the very high and the very low ends of the distribution of scores on all DORF passages were examined. Multiple criteria were considered to locate outlier students: 1) if the scores were at the maximum (or minimum) for the majority of passages - preferentially all passages - and if the student scores skew the distribution of DORF scores; and

2) if any of the scores were greater than the score at the 3rd quartile +1.5* (interquartile range) or less than the score at the 1st quartile -1.5* (interquartile range) (i.e., scores beyond the range of a standard box plot). Additionally, if the student did not have a score on the end-of-year *DIBELS Next* benchmark passage, then the entire record was omitted from analysis.

Individual student data were examined to determine if a highly variable student would compromise conclusions about passage functioning. Very little anecdotal information was provided by examiners. Thus, to explore the variability issue, a regression analysis was performed for each student predicting the student-level mean score with testing session as the dependent variable. For DORF, the regression analysis was also performed by passage type (expository or narrative). The root mean squared error (RMSE) of the prediction was used to flag students whose scores were highly variable. For those students flagged by their RMSE, the individual student scores were evaluated to check for remarkable session-to-session differences in reading performance that did not appear to be related to differences in passage difficulty. Whenever possible, the erroneous scores were set to missing in the data, but in extreme cases, an entire student record was omitted from analysis.

In first grade, three students were omitted from analysis, two students with high scores and one student with low scores. In second grade, two students were omitted from analysis, one student with high scores and one student with low scores. In third grade, two students were omitted from analysis, one student with high scores and one student with low scores. In sixth grade, three students were omitted from analysis, two students with high scores and one student with low scores. Additionally, two students in first grade and three students in second grade were missing one or more of their end-of-year *DIBELS Next* benchmark passages, and

were subsequently omitted from analysis. One student earned an extremely high score for one particular passage, which was inconsistent with the student's record. That score was set to missing for analysis. Some students skipped a line while reading, resulting in an inflated error rate. There were two students in first grade, three students in second grade, and one student in third grade that skipped a line. These students were included in analysis.

Daze. Two scores are recorded for the Daze measure: a score that indicates the number of words correctly chosen and a score that indicates the number of words that were incorrectly chosen to complete the sentence. The sum of these two scores should not be greater than the number of items in the measure. We did not find any pairs of scores for this measure that violated this rule.

Scores for students performing at the very high and the very low ends of the distribution of scores on all Daze passages were examined. Again, we used multiple criteria to locate outlier students: 1) if the scores were at the maximum (or minimum) for the majority of passages - preferentially all passages - and if the student scores skew the distribution of Daze Adjusted Scores; and 2) if the scores were greater than the score at the 3rd quartile +1.5* (interquartile range) (i.e., they are beyond the range of a standard box plot).

Using these procedures, we found two outlier students in third grade; one student had all high scores and one student had all low scores. In addition, two outlier students were found in fourth grade; both students had the two highest scores on all passages. Three outlier students were identified in sixth grade; two students had high scores and one student had low scores.

Daze Guessing. During the initial stages of analysis, some passages were noted as having an inflated error rate. If the error rate was approximately 1/3 of the student's score, the student

booklet was checked for patterns of guessing on all of their Daze passages. We developed the following criteria for flagging instances of suspected guessing on a student-by-student basis: 1) if the scores and errors rates were not consistent with other passages from the same student; and 2) if certain testing sessions had different patterns of reading behavior than others (e.g., if during the first few testing sessions, the student completed one page of the story before time expired and recorded low errors, but for the final testing sessions, the student completed two or more pages of the story before time expired and recorded high errors). If any passage met the above criteria, we flagged the suspected passages for guessing. Additionally, if the end-of-year *DIBELS Next* benchmark passage was flagged by these criteria, then the entire record was omitted from analysis. This procedure was followed because subsequent analysis depended on the student's *DIBELS Next* benchmark passage score.

In third grade, two students were omitted from analysis using the criteria we developed; the first student appeared to have guessed on the *DIBELS Next* passage, and the second student guessed on two of the five *DIBELS AD* passages and did not attempt one of the *DIBELS AD* passages. In fourth grade, one student was omitted from analysis for guessing on the *DIBELS Next* passage. Also in fourth grade, three students each had one individual passage score set to missing for suspected guessing. In fifth grade, three students were omitted from analysis; the first student did not complete the *DIBELS Next* passage, the second student guessed on the *DIBELS Next* passage, and the third student's booklet did not include the first page of the *DIBELS Next* passage and thus their score was incomplete. In sixth grade, two students were omitted from analysis; the first student guessed on the *DIBELS Next* passage, and the second student guessed on three of the five *DIBELS AD* passages.

Daze Adjusted Score. The number of Daze correct responses was adjusted to more accurately reflect the students' overall performance by factoring in the number of incorrect responses. This adjustment was completed by subtracting half the number of incorrect responses from the number of correct responses.

Data analysis. The data analysis included two parts. The first part of the analysis primarily addressed research questions one through five and examined which passages should be included in *DIBELS AD*. For both DORF and Daze, main effects of passage order and passage type (*DIBELS AD* or *DIBELS Next* passages) were checked. The results were evaluated by descriptive statistics and correlational relationships. Next, similarities and differences were explored using the end-of-year *DIBELS Next* benchmark Daze passage as the control passage. The following analyses were performed: analysis of means, control and variability charts, Tukey's Honest Significant Difference (HSD) multiple comparison procedure (which tests all possible pair-wise combinations of passages to see if at least one mean difference is significantly different than zero), nonparametric tests about the median, and tests using the Hsu-Dunnnett method for multiple comparisons to compare all *DIBELS AD* passages to the *DIBELS Next* Median passage. Finally, those passages which were the most similar to the *DIBELS Next* benchmark passages were identified at each grade level.

The second part of the analysis addressed research question six and explored how the different DORF passages should be grouped together into triads, and which Daze passage would be assigned to each DORF passage triad.

DIBELS Oral Reading Fluency (DORF). Information was collected on 13 new DORF passages and compared to the end-of-year *DIBELS Next* DORF benchmark triad. In first, second, and third

grade, information was collected on five expository and eight narrative passages. In fourth, fifth, and sixth grades, information was collected on eight expository and five narrative passages. For first, second, and third grade, each of the three final triads of DORF passages contains one expository and two narrative passages. For fourth, fifth, and sixth grade, each of the three final triads of DORF passages contains two expository and one narrative passage. We also collected information about the session and the order in which the passage was given for each individual passage.

For each grade, narrative and expository passages were first compared to the *DIBELS Next* student-level median separately. Multiple comparison procedures and tests for differences between all passages were performed, and candidates for exclusion were identified in each group. Next, multiple comparison procedures and tests for differences were re-run to evaluate how the remaining passages related to each other. Then, all passages were evaluated together for similarities to the *DIBELS Next* student-level median. For each grade, the nine passages that were most similar to the *DIBELS Next* student-level median were selected for *DIBELS AD*.

Daze. Information was collected on five new Daze passages and the end-of-year *DIBELS Next* benchmark Daze passage. For each individual passage, we also collected information about the session and the order in which the passage was given. For each grade, passages were compared to the end-of-year *DIBELS Next* Daze benchmark passage. Multiple comparison procedures and tests for differences between all passages were performed, and candidates for exclusion were identified. Next, multiple comparison procedures and tests for differences were re-run to evaluate how the remaining passages related to each other. Then, all passages were evaluated together for similarities to the *DIBELS Next* benchmark Daze passage. For each grade,

the three Daze passages that were most similar to the *DIBELS Next* benchmark Daze passage were selected for *DIBELS AD*.

Decision rules for passage selection. Passage-level means, standard deviations, distributions, and relative difficulty (student-level residuals) were compared to the end-of-year *DIBELS Next* benchmark passage(s) to look for similarities and differences. In most cases, passages that were markedly different were quickly identified and excluded. In some cases, there were no stand-out passages to exclude. In those cases, passage combinations that were the most similar were selected. In other cases, more significantly different passages existed than we could exclude. In those cases, we selected passage combinations that were the least significantly different.

Results

This section presents the results of analyses examining the *DIBELS AD* Oral Reading Fluency and Daze passages relationship to the end-of-year *DIBELS Next* benchmark test materials. Sample size descriptions for the various subsets of data are described, followed by descriptive statistics and reliability information for the *DIBELS AD* passages. Following the reliability data, we present data on passage selection for inclusion/exclusion. Finally, we present data for the formation of three-passage triads for the secure version of spring/end-of-year benchmark assessment.

Descriptive Statistics and Passage Selection Analyses

Research questions one through three were addressed first: 1) what are the mean DORF and Daze passages at each grade level, 2) what are the most extreme passages at each grade level relative to the mean passage, and 3) how do the new passage scores compare to scores on

current DIBELS Next spring benchmark passages? These research questions were evaluated by examining the descriptive statistics (mean, standard deviation, percentiles, and accuracy) for each DORF passage by grade (see Table 3) and for each Daze passage by grade (see Table 4). Scores are reported for each proposed *DIBELS AD* passage and the end-of-year *DIBELS Next* benchmark passage. Passages are numbered according to grade; the number preceding the decimal indicates the grade level.

Preliminary Analysis. Passages were evaluated for each student by examining scatter plots of passages' scores across testing sessions (see Figure 1). Descriptive statistics and correlations were evaluated for consistency. Tests were performed for all main effects, and the results are summarized in Table 5. For both DORF and Daze, the first passage given in the first session almost always resulted in a significantly lower score than the student average.

DIBELS Oral Reading Fluency. Effect tests for differences in DORF passages were significant for fourth, fifth, and sixth grades (see Table 5) indicating that at least one pair of passage mean scores was significantly different in each of those grades. Effect tests for passage type (AD or *DIBELS Next*) were not significant. Fourth and fifth grade reported significant differences between the expository and narrative group means.

In fifth grade, we found significant order effects. There was a straight-line linear relationship between the order in which the passages were given and the number of DORF words read correctly (WRC). The slope of the effect was .69 ($p = .007$). The order effect was removed by calculating the residuals from a simple linear regression predicting DORF WRC with order as the dependent variable. The residuals from this regression analysis were used in place of DORF WRC for all subsequent tests. This procedure removed the order effects while preserving

individual passage information relative to every other passage and the *DIBELS Next* benchmark passages.

The new passages are numbered by grade from 1-13, and the end-of-year *DIBELS Next* DORF benchmark triad passages are numbered 14 – 16. The *DIBELS Next* student-level median score is labeled with an “M” and was calculated for each student from the end-of-year *DIBELS Next* DORF benchmark triad (i.e., passages 14, 15, and 16). For example, passage 1.5 refers to grade 1 passage 5, and passage 1.M refers to the grade 1 *DIBELS Next* median passage (the median of passages 14, 15, and 16).

First grade. There was suggestive, but inconclusive, evidence indicating a difference between the mean of the *DIBELS Next* median passage 1.M and passages 1.2 ($p = .05$) and 1.12 ($p = .09$). There was no evidence to suggest differences between passage 1.M and every other passage, narrative or expository. There were no pair-wise differences using Tukey’s HSD, and so comparisons using Student’s *t* procedure provided evidence for exclusion. Among the passages, some evidence indicated differences between narrative passage 1.2 and narrative passages 1.1, ($p = .03$), 1.7 ($p = .007$), and 1.8 ($p = .02$), and expository passages 1.11 and 1.12 ($p = .02$). Evidence also indicated differences between narrative passages 1.4 and 1.7 ($p = .03$). The magnitude of the differences and p -values suggested that narrative passages 1.2 and 1.7 and expository passages 1.11 and 1.12 were candidates for exclusion.

When the exclusion candidates were removed, tests for differences in passage-level mean scores returned non-significant results. The passage-level residuals for the remaining passages indicated significant differences in relative difficulty between passage 1.M and narrative passages 1.3 ($p = .04$), 1.4 ($p < .001$), 1.5 ($p = .003$), and 1.6 ($p = .02$), and for all three remaining

expository passages, 1.9, 1.10, and 1.13 ($p < .005$ for all). The final passages chosen for first-grade DORF were narrative passages 1.1, 1.3, 1.4, 1.5, 1.6, and 1.8, and expository passages 1.9, 1.10, and 1.13.

Second grade. We found no evidence to suggest a difference between passage 2.M and every other passage, narrative or expository. No pair-wise differences were found using Tukey's HSD, and so comparisons using Student's t procedure provided evidence for exclusion. Among the passages, there was suggestive, but inconclusive, evidence for differences between passage 2.M and narrative passage 2.4 ($p = .05$) and 2.5 ($p = .05$), and no evidence for differences between expository passages. Though not significant, the magnitude of the differences (approximately 20 words correct) and p -values indicated that narrative passages 2.4 and 2.5 were candidates for exclusion. The expository passages selected as candidates for inclusion were those passages whose mean scores and relative difficulty were most similar to the *DIBELS Next* Student-Level Median, which meant excluding passages 2.10 and 2.13.

When the exclusion candidates were removed, tests for differences in passage-level mean score returned non-significant results. The passage-level residuals for all of the remaining passages together indicated significant differences in relative difficulty between passage 2.M and narrative passages 2.1 ($p < .005$), 2.3 ($p < .001$), 2.7 ($p < .005$), and 2.8 ($p = .01$). The final passages selected for second-grade DORF were narrative passages 2.1, 2.2, 2.3, 2.6, 2.7, and 2.8, and expository passages 2.9, 2.11, and 2.12.

Third grade. Similar to second grade, we found no evidence to suggest a difference between passage 3.M and every other third-grade passage, narrative or expository. No pair-wise differences were found using Tukey's HSD, and so comparisons using Student's t procedure

provided evidence for exclusion. Among the passages, some evidence indicated differences between narrative passages 3.4 and 3.3 ($p = .03$), and narrative passages 3.4 and 3.7 ($p = .03$), and for differences between expository passage 3.12 and 3.9 ($p = .02$) and 3.12 and 3.M ($p = .03$). The magnitude of the differences (approximately 20 Words Correct) and p -values indicated that narrative passages 3.3 and 3.7 and expository passage 3.12 were candidates for exclusion. The final expository passages that were selected as candidates for inclusion were those passages whose mean scores and relative difficulty were most similar to the *DIBELS Next* Student-Level Median, which meant also excluding passage 3.11.

When the exclusion candidates were removed, tests for differences in passage-level mean score returned non-significant results. The passage-level residuals for all of the remaining passages together indicated significant differences in relative difficulty between 3.M and narrative passages 3.2 ($p = .003$), 3.4 ($p = .008$), 3.8 ($p = .005$), and 3.10 ($p < .001$). The final passages selected for third-grade DORF were narrative passages 3.1, 3.2, 3.4, 3.5, 3.6, and 3.8, and expository passages 3.9, 3.10, and 3.13.

Fourth grade. At this grade level, we found suggestive, but inconclusive, evidence indicating a difference between passage 4.M and expository passage 4.8 ($p = .007$). We did not find evidence to suggest a difference between passage 4.M and any of the other passages, narrative or expository. No pair-wise differences were found using Tukey's HSD, and so comparisons using Student's t procedure provided evidence for exclusion. Among the passages, some evidence indicated differences between narrative passages 4.2 and 4.7 ($p = .04$), convincing evidence was found for differences between expository passages 4.8 and 4.12 ($p = .006$), 4.8 and 4.13 ($p = .003$), and some evidence for differences in expository passages 4.4 and 4.8 ($p =$

.01), 4.8 and 4.10 ($p = .03$), and 4.11 and 4.M ($p = .02$). The magnitude of the differences (approximately 20 Words Correct) and p -values suggested that narrative passages 4.2 and 4.7 and expository passage 4.8 and 4.11 were candidates for exclusion.

When the exclusion candidates were removed, tests for differences in passage-level mean score returned non-significant results. The passage-level residuals for all of the remaining passages together indicated significant differences in relative difficulty between passage 4.M and expository passages 4.5 ($p < .001$), 4.9 ($p < .001$), and 4.10 ($p = .01$). The final passages selected for fourth-grade DORF were narrative passages 4.1, 4.3, 4.6, and expository passages 4.4, 4.5, 4.9, 4.10, 4.12, and 4.13.

Fifth grade. We found convincing evidence to suggest a difference between passage 5.M and expository passage 5.9 ($p = .002$). In addition, we found suggestive, but inconclusive, evidence for differences between 5.M and expository passages 5.10 ($p = .05$) and 5.11 ($p = .06$). Finally, there was some evidence to suggest a difference between passage 5.M and narrative passage 5.5 ($p = .02$). Pair-wise comparisons using Tukey's HSD among the passages provided overwhelming evidence for differences between narrative passages 5.2 and 5.5 ($p < .001$), convincing evidence for differences between narrative passages 5.1 and 5.5 ($p = .003$), 5.3 and 5.5 ($p = .008$), and some evidence for differences between narrative passages 5.4 and 5.5 ($p = .04$). Also, we found suggestive, but inconclusive, evidence indicating differences between 5.M and 5.5 ($p = .06$). Among expository passages, Tukey's HSD provided overwhelming evidence indicating differences between passages 5.6 and 5.9 ($p < .001$), 5.7 and 5.9 ($p = .001$), and 5.M and 5.9 ($p = .006$). There was some evidence for differences between 5.8 and 5.9 ($p = .02$), and between 5.6 and 5.10 ($p = .02$) and 5.6 and 5.11 ($p = .02$). The magnitude of the differences and

p -values with the abundance of significant differences provided clear evidence to exclude narrative passage 5.5 and expository passage 5.9. The final expository passages selected as candidates for inclusion were those passages whose mean scores and relative difficulty were the closest to the *DIBELS Next* Student-Level Median, which meant also excluding narrative passage 5.2 and expository passage 5.10.

When the exclusion candidates were removed, tests for differences in passage-level mean score returned non-significant results. The passage-level residuals for all of the remaining passages together indicated significant differences in relative difficulty between passage 5.M and expository passages 5.11 ($p < .001$) and 5.12 ($p = .005$). The final passages selected for fifth-grade DORF were narrative passages 5.1, 5.3, and 5.4, and expository passages 5.6, 5.7, 5.8, 5.11, 5.12, and 5.13.

Sixth grade. We found no evidence to suggest a difference between 6.M and every other passage, narrative or expository. In addition, there were no pair-wise differences using Tukey's HSD, and so comparisons using Student's t procedure provided evidence for exclusion. Among the narrative passages, there was overwhelming evidence for differences between passages 6.4 and 6.5 ($p < .001$), convincing evidence for differences between passages 6.2 and 6.4 ($p = .004$), 6.1 and 6.4 ($p < .004$), and some evidence for differences between 6.M and passage 6.4 ($p = .02$), and passages 6.3 and 6.5 ($p = .02$). Among the expository passages, we found some evidence for differences between passages 6.7 and 6.11 ($p = .02$), 6.7 and 6.13 ($p = .03$), 6.9 and 6.11 ($p = .03$), 6.9 and 6.13 ($p = .03$), 6.11 and 6.12 ($p = .05$). The magnitude of the differences and p -values with the abundance of significant differences provided clear evidence to exclude narrative passage 6.4. The final passages selected as candidates for inclusion were those

passages whose mean scores and relative difficulty were most similar to the *DIBELS Next* Student-Level Median, which meant also excluding narrative passage 6.5 and expository passages 6.7 and 6.9.

When the exclusion candidates were removed, tests for differences in passage-level mean score returned non-significant results. The passage-level residuals for all of the remaining passages together also indicated no significant differences in relative difficulty between passage 6.M and all other passages, expository or narrative. The final passages selected for sixth-grade DORF are narrative passages 6.1, 6.2, and 6.3, and expository passages 6.6, 6.8, 6.10, 6.11, 6.12, and 6.13.

Daze. Effect tests for differences in Daze passages were significant for fourth, fifth, and sixth grades (see Table 5) indicating that at least one pair of passage mean scores was significantly different in each of those grades. We found significant differences between the *DIBELS AD* passages and the end-of-year *DIBELS Next* benchmark passage in third and fourth grades ($p < .001$ on both). Students performed better on the end-of-year *DIBELS Next* benchmark passage than the new *DIBELS AD* passages, scoring approximately four points higher on average.

In addition, we found significant order effects in fourth, fifth, and sixth grades. There was a straight-line linear relationship between the order in which the passages were given and the Daze Adjusted Score (and the Daze raw score). The slope of the relationship was steeper as grade level increased; fourth-grade slope was .70 ($p = .01$), fifth-grade slope was 1.02 ($p < .001$), and sixth-grade slope was 1.18 ($p < .001$). For these grades, we removed the order effect with the same procedure used for DORF Words Correct, by calculating the residuals from a simple linear regression predicting the Daze Adjusted Score with order as the dependent variable. The

residuals from this regression analysis were used in place of the Daze Adjusted Score for all subsequent tests. Using this procedure, we removed the order effects while preserving individual passage information relative to every other passage and the *DIBELS Next* benchmark passage. Passages are numbered from 1-5, and the *DIBELS Next* benchmark Daze passage is passage 6.

Third grade. We did not find any evidence to suggest a difference between passage 3.6 and passages 3.2 and 3.3 ($p = .27$, and $.18$, respectively). However, we found evidence for a difference between passages 3.1 and 3.6 ($p = .02$), indicating that the most difficult passage of the five was the only passage that was significantly different from passage 3.6.

To select the third and final passage, the remaining passages, 3.4 and 3.5, were evaluated for their similarities to passages 3.2 and 3.4. Variability control charts with control limits for the mean passage-level residual were evaluated for the two separate groups of passages; group 1 contained passages 3.2, 3.3, and 3.4 grouped with passage 3.6, and group 2 contained passages 3.2, 3.3, and 3.5 grouped with passage 3.6. The results indicated that the relative difficulty between passages was less varied within group 1. Therefore, the passages chosen for third-grade Daze were passages 3.2, 3.3, and 3.4.

Fourth grade. We found overwhelming evidence to suggest significant differences in mean score between passage 4.6 and passages 4.2, 4.3, and 4.5 ($p < .001$, each), convincing evidence for differences between passage 4.6 and 4.4 ($p = .02$), and suggestive evidence for a difference with passage 4.1 ($p = .05$). Tukey's multiple comparison procedure and nonparametric tests using Dunnett's method for joint ranking suggested that passages 4.1 and 4.4 were not significantly different from passage 4.6. Looking at the magnitude of the effect test statistics

and p-values, passages 4.1 and 4.4 were the least significantly different of all the passages when compared to 4.6, the *DIBELS Next* benchmark passage. Therefore, these two passages were selected for inclusion.

To choose the third and final passage, the remaining passages, 4.2, 4.3, and 4.5, were evaluated for their similarities to passages 4.1 and 4.4. Tukey's multiple comparison procedure indicated that passage 4.5, the most difficult of the five passages, was significantly different from 4.1 ($p < .001$), and thus, we eliminated it from further evaluation. Variability control charts with control limits for the mean for the passage-level residual were evaluated for the two separate groups of passages; group 1 contained passages 4.1, 4.2, and 4.4 grouped with passage 4.6, and group 2 contained passages 4.1, 4.3, and 4.4 grouped with passage 4.6. The results indicated that the relative difficulty between passages was less varied within group 1. Therefore, the passages selected for fourth-grade Daze were passages 4.1, 4.2, and 4.4.

Fifth grade. We found a significant difference between passage 5.6 and the most difficult passage, 5.3 ($p = .006$). Thus, we eliminated passage 5.3 from further evaluation. Tukey's multiple comparison procedure indicated that passages 5.1 and 5.4 were not significantly different from each other nor were either passage significantly different from passage 5.6. Therefore, passages 5.1 and 5.4 were set for inclusion.

To choose the third and final passage, the remaining two passages, 5.2 and 5.5, were evaluated for their similarities to passages 5.1, 5.4, and 5.6. Variability control charts with control limits for the mean passage-level residual were evaluated for the two separate groups of passages; group 1 contained passages 5.1, 5.2, and 5.4 grouped with passage 5.6, and group 2 contained passages 5.1, 5.4, and 5.5 grouped with passage 5.6. The results indicated passages

were less varied within group 2; all passages fell within the control limits. Therefore, the passages selected for fifth-grade Daze were passages 5.1, 5.4, and 5.5.

Sixth grade. We found no evidence to suggest differences between passage 6.6 and any other passage. Evaluating all pair-wise comparisons using Tukey's multiple comparison procedure and Wilcoxon's Rank-Sum method provided evidence of differences between passage 6.5 and passages 6.2, 6.3, and 6.4 ($p < .005$ for all), and evidence of differences between passage 6.1 and passages 6.2, 6.3, and 6.4 ($p < .01$ for all).

To choose a group of three passages, variability control charts with control limits for the mean for the passage-level residual were examined for several groups of passage combinations. The passage combination that contained the least between-passage variability was the group consisting of passages 6.2, 6.3, and 6.4 grouped with passage 6.6. All means fell within control limits. Therefore, the passages selected for sixth-grade Daze were passages 6.2, 6.3, and 6.4.

Characteristics and Features of Outlier/Discarded Passages

To address research question four regarding outlier passage characteristics, we examined the DMG passage difficulty indices, along with any notes made by assessors on the passages. Tables A1 - A10 in the appendix contain all of the data from the DMG Passage Revision Utility, including the DMG Passage Difficulty Index. Discarded passages are noted with an asterisk. Equal numbers of expository and narrative passages were discarded at each grade level. Furthermore, the data in Tables A1 - A10 do not indicate any consistent patterns of differences between discarded passages and the retained passages.

Anecdotal notes on the DORF passages also were examined for patterns. Assessors' notes were made with respect to: Challenging Words, Topic (impact on fluency, appropriate for

grade, and/or interest to student), Difficult Proper Names, and Difficult Sentence Structure.

First, all rows in the data set that contained notes about a passage were extracted from the master data set. Next, the notes made on the passages were examined by grade level.

Anecdotal notes were recorded for both AD passages and *DIBELS Next* passages. Not all of the passages had notes and only the passages with notes could be compared. For the passages with notes, the average number of notes made for each category was examined and compared across all passages at each grade level. Nineteen discarded passages had notes, and 40 retained passages had notes. In order to compare discarded and retained passages, the mean number of notes per passage type was examined for each category. Small differences between discarded and retained passages were found of approximately 1 note more for discarded passages, on average, across all four categories.

Alternate-Form Reliability

We examined alternate form reliability to address research question five. Alternate-form reliability is a means of generalizing to different item samples. Students are tested with different, but "equivalent," forms of the test and scores from these two tests are correlated. Correlational strength is based on criteria delineated by Hopkins (2002). Alternate-form reliability data for DORF are reported by grade in Tables 6 through 11. Alternate-form reliability for Daze is reported in Table 12. Correlations were evaluated for the magnitude of the correlation and the significance with the *DIBELS Next* benchmark passage as well as the other *DIBELS AD* passages.

DIBELS Oral Reading Fluency. With few exceptions, the results of our analyses indicated strong to nearly perfect correlations between all DORF passages. Correlations ranged from .87

to .97 in first grade, .83 to .99 in second grade, .77 to .95 in third grade, .67 to .97 in fourth grade, .59 to .95 in fifth grade, and .49 to .98 in sixth grade.

Daze. With some exceptions, our analyses yielded correlations between all Daze passages that were moderate-strong to strong. Correlations ranged from .59 to .78 in third grade, .46 to .65 in fourth grade, .44 to .74 in fifth grade, and .56 to .73 in sixth grade.

Triad Formation

To address research question six, we sought to organize the nine selected DORF passages into three groups of three (three triads). Each triad was then paired with a single Daze passage. Each of the final three sets of passages are intended to be administered securely at the end of the school year across three consecutive years (i.e., one set of passages per year). The main statistic used in this process was the passage-level residual, which is an estimate of difficulty relative to all other passages within each grade. We calculated the passage-level residual by first performing a regression analysis for each student predicting the student's score (adjusted for order effects when appropriate) using session as the dependent variable. The residual for that passage at the student-level is the difference between the predicted session score and the actual individual passage score (see Figure 1). Next, for each passage, the student-level residuals were summed across all students and divided by the total number of students to calculate the mean of all of the student-level residuals. This mean score was the passage-level residual. The same process was repeated for each passage for both DORF and Daze. Passage-level residuals for DORF are reported in Table 13, and for Daze in Table 14.

Triads were formed from two narrative passages and one expository passage in first, second, and third grades, and two expository passages and one narrative passage in fourth,

fifth, and sixth grades. All possible triads with these combinations were examined. For each possible triad, the passage-level residuals were summed. The absolute value of the sum was calculated. For ease, we will call this statistic $|\Sigma res|$. For each grade, a threshold was identified for $|\Sigma res|$ that included enough triads so that all nine passages were considered. Next, we counted the occurrence of each passage within a triad. The first triad was selected based on low occurrences. Once the first triad was chosen, other triads that contained any of the three first-triad passages were eliminated. The second triad was selected by the same process. The third triad was formed from the remaining three passages as long as $|\Sigma res|$ fell within the threshold.

To illustrate this process, consider first grade as an example. A threshold of 2 was identified for triad formation. There were nine triads in which $|\Sigma res| < 2$. The lowest occurring passages were 1.6 and 1.5, both of which appeared in only two of the nine possible triad combinations. Randomly, passage 1.6 was selected first. The triad containing passage 1.6 with the lowest $|\Sigma res|$ was chosen, and contained passages 1.1, 1.6, and 1.9, $|\Sigma res| = 1.51$. This triad was later discarded, because mutually exclusive triads could not be formed from the remaining passages where $|\Sigma res| < 2$. Therefore, the other triad containing passage 1.6 was selected, and contained passages 1.3, 1.6, 1.10, $|\Sigma res| = 1.95$. Passage 1.5, which appeared in two out of the remaining eight triads, was then selected. The triad with the smallest $|\Sigma res|$ was again chosen, which contained passages 1.1, 1.5, and 1.9, $|\Sigma res| = 0.3$. The third and final triad was formed from the remaining three passages, 1.4, 1.8, and 1.13, $|\Sigma res| = 1.63$, which was less than 2.

Once formed, triads for grades three through six were paired with one of the selected Daze passages. The Daze passages were chosen based on the strength of the passage-level residual. The DORF triad that had the smallest passage-level residual (i.e., the most difficult triad), was paired with the Daze passage that had the largest passage-level residual (i.e., the easiest Daze passage). The middle triad was paired with the middle Daze passage, and the most difficult triad was paired with the easiest Daze passage.

The passages in each triad, the $|\Sigma res|$ for each triad, and the paired Daze passage is reported in Table 15. Correlations between triads and the *DIBELS Next* end-of-year benchmark triad is reported in Table 16. Matched-pairs *t*-tests were performed to evaluate mean differences between the newly formed DORF triads and the *DIBELS Next* end-of-year DORF benchmark triad. Results are reported in Table 17. And finally, matched-pairs *t*-tests were performed to evaluate mean differences between the Daze passages and the *DIBELS Next* end-of-year Daze benchmark passage. Results are reported in Table 18.

Discussion

In this study, we used a multi-method approach to understand passage variability and how to assign reading passages for use in a secure version of *DIBELS* (*DIBELS* for accountability decisions, *DIBELS AD*). We used the same passage selection process for each grade. This process examined differences between the newly created DORF and Daze passages and passages used for end-of-year benchmark assessment with *DIBELS Next* (both DORF and Daze). This process resulted in the selection of nine DORF passages per grade in grades one through six and three Daze passages per grade in grades three through six. Passages were selected to minimize differences from *DIBELS Next* and our results support that these differences were minimized. In

addition, minimal qualitative differences were noted between the discarded and retained passages from the newly created set of passages. Furthermore, we found exceptionally strong alternate form reliability for the retained passages, at the individual DORF and Daze passage level. DORF passage triads were formed in such a way as to further minimize differences between *DIBELS Next* and *DIBELS AD* materials. Finally, *DIBELS AD* Daze passages were matched to *DIBELS AD* DORF triads in such a way as to attempt to balance any differences in difficulty between *DIBELS Next* and *DIBELS AD*.

Limitations

As with every study, ours had some limitations. One limitation of this study is the relatively small sample examined for our data analysis. While adequate for the analyses we conducted, the sample size we had does limit the external validity/generalizability of the results. In addition, we did not have student-level demographic data available to us on the participants in this study. Thus, we examined demographic data at the school level. The degree to which the sample of students we tested is, or is not, representative of the school demographics is unknown to us. Thus, making statements about sample representativeness with respect to other students who may be tested with these materials in the future more tenuous. Finally, although training on administration and scoring of the measures and fidelity to procedures was checked at the beginning of the study, score re-checking was not checked during data collection.

Implications & Conclusion

Despite our efforts to reduce differences, some of the newly created *DIBELS AD* passages appeared to be more difficult overall than the *DIBELS Next* passages. We attempted to balance

these differences during triad formation and when matching DORF passage triads to a Daze passage. However, it is possible that the newly created reading passages function differently enough from *DIBELS Next* in practice that more formal equating procedures should be conducted. This concern is particularly relevant given the high-stakes nature of *DIBELS AD* (i.e., where student reading data may contribute to accountability decisions, including teacher evaluation). Therefore, future research will include equating studies for *DIBELS AD* designated materials to further reduce any concerns about differences between *DIBELS Next* and *DIBELS AD* results.

References

- Ardoin, S. P., Suldo, S. M., Witt, J., Aldrich, S., & McDonald, E. (2005). Accuracy of readability estimates' predictions of CBM performance. *School Psychology Quarterly, 20*(1), 1-22. doi: [10.1521/scpq.20.1.1.64193](https://doi.org/10.1521/scpq.20.1.1.64193)
- Betts, J., Pickart, M., & Heistad, D. (2009). An investigation of the psychometric evidence of CBM-R passage equivalence: Utility of readability statistics and equating for alternate forms. *Journal of School Psychology, 47*(1), 1-17. DOI: 10.1016/j.jsp.2008.09.001. ISSN: 0022-4405.
- Compton, D. L., Appleton, A. C., & Hosp, M. K. (2004). Exploring the relationship between text-leveling systems and reading accuracy and fluency in second grade students who are average and poor decoders. *Learning Disabilities Research and Practice, 19*(3), 176-184. DOI: 10.1111/j.1540-5826.2004.00102.x
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232.
- 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*(3) 315-342. doi: 10.1016/j.jsp.2007.06.003.
- Good, R. H., III, & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 61 - 88). New York: Guilford.
- Good, R. H., Kaminski, R. A., Dewey, E. N., Wallin, J., Powell-Smith, K. A., & Latimer, R. J. (2013). *DIBELS Next Technical Manual*. Eugene, OR: Dynamic Measurement Group. Available: <http://dibels.org/next>

- Hintze, J. M., & Christ, T. J. (2004). An examination of variability as a function of passage variance in CBM progress monitoring. *School Psychology Review, 33*(2), 204-217.
- Hopkins, W. G. (2002). A scale of magnitudes for the effect statistics. In *A Review of Statistics*. Retrieved July 30, 2010 from <http://www.sportsci.org/resource/stats/effectmag.html>
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Reports of the subgroups. Washington, DC: National Institute of Child Health and Human Development.
- Powell-Smith, K. A., & Bradley-Klug, K. L. (2001). Another look at the "C" in CBM: Does it really matter if curriculum-based measurement reading probes are "curriculum-based?" *Psychology in the Schools, 38* (4), 299-312.
- Powell-Smith, K. A., Good, R. H., III, & Atkins, T. (2010). *DIBELS Next Oral Reading Fluency readability study*. Technical Report No. 7, Dynamic Measurement Group, Eugene, Oregon.
- Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford.
- Tindal, G., Marston, D., & Deno, S. L. (1983). *The reliability of direct and repeated measurement* (Research Rep. 109). Minneapolis, MN: University of Minnesota Institute for Research on Learning Disabilities.
- United States Census Bureau (n.d.). http://www.census.gov/geo/maps-data/maps/pdfs/reference/us_regdiv.pdf. Retrieved January 22, 2014.
- Williams, K. T. (2001). *Group Reading and Diagnostic Evaluation (GRADE)*. New York: Pearson.

Table 1

School Demographic Characteristics

	School Number	
	1	2 ^a
Locale	Rural: Remote	Suburb: Large
Grades Taught	KG - 6	PK - 8
Total Students	392	731
Student/Teacher Ratio	23:1	19:1 ^a
Title 1 Eligible	Yes	--
Free/Reduced Price Lunch	54%	--
Percent Female	49%	--
Student Ethnicity/Race		
American Indian/Alaska Native	13%	0%
Asian/Pacific Islander	< 1%	2%
Black	< 1%	< 1%
Hispanic	0%	1%
White	86%	96%
Two or More Races	0%	0%

Note. Demographic characteristics gathered at the school level from the National Center for Education Statistics website: nces.ed.gov. Data for school 1 represent the student population from the 2010-2011 school year, and data from school 2 represents the student population from the 2011-2012 school year.

^a Student/Teacher ratio and ethnicity data are reported for K - 8 grade levels only (n = 678).

Table 2

Percent of the Sample At or Above the Winter and Spring DIBELS Next Benchmark Goals for DORF and Daze

Grade	Winter Benchmark Goal	% At or Above Winter Goal	Spring Benchmark Goal	% At or Above Spring Goal
<i>DORF</i>				
First	23	88.46	47	61.54
Second	72	81.48	87	66.67
Third	86	85.71	100	67.86
Fourth	103	92.59	115	70.37
Fifth	120	80.00	130	60.00
Sixth	109	92.31	120	80.77
<i>Daze</i>				
Third	11	100.00	19	85.00
Fourth	17	100.00	24	92.31
Fifth	20	92.59	24	77.78
Sixth	19	100.00	21	88.00

Note. Sample sizes: DORF = 159, Daze = 103. Percent signs omitted. Testing occurred in February and March.

Table 3

Descriptive Statistics for Oral Reading Fluency Passages for All Grades

Passage Number	Passage Type	Title	N	WRC		Percentiles					Accuracy	
				Mean	SD	5th	25th	50th	75th	95th	Mean	SD
<i>First Grade</i>												
1.1	AD Nar	Visiting Farm	25	62.52	27.50	26.60	37.00	64.00	79.50	117.40	85.50	14.46
1.2*	AD Nar	Catching Frogs	25	64.60	27.82	26.20	40.50	60.00	84.50	112.40	93.43	6.88
1.3	AD Nar	A Trip Through the Car Wash	26	70.08	30.18	28.50	43.00	69.50	86.00	136.40	91.51	7.73
1.4	AD Nar	What's in the Box?	26	61.19	35.29	15.40	26.00	59.50	89.00	125.15	92.06	8.45
1.5	AD Nar	The Crows Are in the Corn	26	51.19	27.40	15.05	24.25	50.50	69.25	109.40	91.72	8.00
1.6	AD Nar	In the Night Sky	26	52.58	32.16	11.40	17.75	56.00	73.25	117.65	91.44	9.87
1.7*	AD Nar	A Thrilling Ride	26	49.62	26.10	13.45	27.50	52.00	64.00	109.35	85.71	13.95
1.8	AD Nar	A Muddy Mess	26	61.62	30.39	20.40	37.00	59.00	78.50	118.90	86.74	15.08
1.9	AD Exp	Everywhere We See Bugs	25	73.28	29.43	33.40	46.00	72.00	91.50	129.00	91.99	8.17
1.10	AD Exp	Making Crepes	26	62.27	31.13	19.70	34.50	59.00	86.50	118.30	86.38	11.28
1.11*	AD Exp	Wind and Rain	26	63.38	31.97	27.05	35.00	56.00	85.00	122.25	86.22	12.25

1.12*	AD Exp	Homes, Homes, Homes	26	54.85	33.31	16.05	24.50	53.50	63.00	124.10	93.31	6.77
1.13	AD Exp	A Frog's Life	25	68.28	34.78	16.20	31.50	66.00	95.00	130.50	89.03	9.99
1.14	DN Nar	The Cocoa Stand	26	52.96	25.63	10.45	31.00	51.50	77.25	97.20	89.92	12.83
1.15	DN Exp	Parts of a Tree	26	52.92	26.85	19.35	26.00	52.00	75.00	107.10	88.22	14.51
1.16	DN Nar	Going to Market	26	65.58	33.32	20.00	34.00	67.50	89.25	128.75	88.20	10.49
1.M	DN	<i>DIBELS Next</i> Median	26	56.31	27.61	19.35	27.75	60.00	77.75	107.80	88.15	12.05

Second Grade

2.1	AD Nar	Popsicle Stand	26	101.85	36.03	46.40	77.25	93.50	127.50	173.30	96.36	2.66
2.2	AD Nar	Pen Pals	27	100.00	38.60	41.80	67.00	97.00	137.00	165.20	96.56	3.29
2.3	AD Nar	A New Friend	26	103.27	33.28	51.10	77.75	95.50	137.50	157.25	95.06	4.14
2.4*	AD Nar	Firefly Night	27	94.26	32.73	45.00	60.00	92.00	125.00	148.00	93.81	5.82
2.5*	AD Nar	Bakery Bear	26	91.23	27.26	52.70	71.25	84.00	120.00	140.10	95.25	3.87
2.6	AD Nar	Sailing	26	91.12	35.48	43.35	58.75	86.50	128.25	151.45	98.15	2.58
2.7	AD Nar	Lunchbox Surprise	26	91.50	35.93	39.05	64.00	86.00	122.50	147.65	96.38	3.73
2.8	AD Nar	Grandma's Attic	27	90.15	26.08	47.00	74.00	86.00	116.00	134.20	94.81	5.50
2.9	AD Exp	Grow Your Own Salad	27	96.70	36.03	42.20	64.00	90.00	124.00	156.60	96.46	4.11
2.10*	AD Exp	What Causes the Seasons?	26	93.65	28.14	48.15	72.75	89.50	121.00	138.30	95.48	5.29

2.11	AD Exp	Hermit Crabs and Their Shells	26	102.42	35.16	54.00	75.00	90.00	131.75	173.20	97.17	3.49
2.12	AD Exp	Making Tacos	27	94.59	29.59	49.20	72.00	87.00	130.00	140.20	95.20	5.59
2.13*	AD Exp	Splish Splash!	26	107.92	30.76	57.55	82.50	103.50	132.00	163.75	94.71	5.85
2.14	DN Nar	Gavin's Jump	27	105.59	38.75	56.40	73.00	94.00	143.00	172.60	95.13	6.78
2.15	DN Exp	The New Year	27	106.85	44.22	40.20	68.00	102.00	149.00	179.80	96.12	5.05
2.16	DN Nar	Roller Skating Fun	27	97.70	30.40	55.00	67.00	94.00	117.00	151.80	96.57	3.13
2.M	DN	<i>DIBELS Next</i> Median	27	104.81	38.00	55.00	68.00	96.00	139.00	172.60	96.90	4.43

Third Grade

3.1	AD Nar	Going on a Field Trip to a Nature Preserve	28	112.89	30.27	51.70	89.25	116.00	139.75	155.30	97.35	3.01
3.2	AD Nar	A Doggie Dog World	28	110.89	29.80	61.30	87.75	106.50	126.00	166.95	97.59	2.25
3.3*	AD Nar	Tool Shed Surprise	28	117.00	25.42	70.80	98.50	118.50	141.75	154.50	97.36	3.90
3.4	AD Nar	Little Silva	28	122.18	31.59	61.80	96.00	119.50	151.50	168.05	97.59	2.49
3.5	AD Nar	A New View of Vegetables	28	103.43	23.67	54.30	85.00	105.50	120.75	143.55	97.79	1.83
3.6	AD Nar	The Most Unusual Hats	27	112.44	30.56	54.80	88.00	112.00	143.00	157.00	97.83	2.27
3.7*	AD Nar	Icing on the Cake	28	123.25	33.08	65.40	93.25	122.50	149.50	172.75	95.46	5.78
3.8	AD Nar	Pecos Bill	28	121.14	33.63	55.55	105.00	119.50	144.75	181.40	97.30	2.32
3.9	AD Exp	Kenya Has a Lot to Offer	27	108.44	35.05	50.40	83.00	103.00	140.00	169.80	97.87	2.42

3.10	AD Exp	Africa's Natural Resources	28	127.61	34.89	56.50	101.00	137.00	153.50	182.95	97.72	3.38
3.11*	AD Exp	Geography of Crete	27	112.81	26.99	64.00	92.00	113.00	129.00	160.20	96.47	2.35
3.12*	AD Exp	Sink or Float?	27	119.67	33.55	63.40	90.00	124.00	146.00	177.60	95.70	4.84
3.13	AD Exp	Henry Ford's Cars	28	108.36	33.61	51.00	91.00	102.00	135.75	178.10	98.27	1.51
3.14	DN Nar	A Surprising Discovery	28	110.71	29.33	60.35	84.25	113.00	135.25	156.30	97.31	3.32
3.15	DN Nar	A Day for a Shadow Dance	28	126.75	37.88	53.45	105.00	126.50	151.00	198.10	96.94	3.47
3.16	DN Exp	A Triple Challenge	28	125.68	35.89	58.05	97.25	131.50	156.50	180.00	96.35	2.68
3.M	DN	<i>DIBELS Next Median</i>	28	120.39	33.03	59.25	97.25	125.50	147.50	174.60	97.11	3.08

Fourth Grade

4.1	AD Nar	Just One Wheel?	27	134.33	30.67	78.20	115.00	136.00	153.00	192.60	97.85	1.75
4.2*	AD Nar	Back to Earth	27	136.04	26.31	83.40	116.00	135.00	157.00	183.00	98.65	1.25
4.3	AD Nar	A History Lesson	27	131.93	20.19	88.40	118.00	132.00	145.00	171.00	97.77	2.79
4.4	AD Exp	Plants Have Needs, Too	26	127.85	29.14	76.70	103.75	129.00	146.25	185.35	99.39	0.76
4.5	AD Exp	Stages of Life	26	114.81	19.55	86.00	97.25	112.00	131.50	146.85	98.35	1.45
4.6	AD Nar	Quite A Game	27	126.48	25.89	70.80	112.00	132.00	143.00	173.00	97.74	2.78
4.7*	AD Nar	Cat in the Box	26	122.42	26.44	70.70	106.50	125.00	142.00	164.95	98.37	1.42
4.8*	AD Exp	Telescopes	27	130.15	23.35	95.00	111.00	129.00	147.00	170.40	97.61	1.49

4.9	AD Exp	The Bill of Rights	26	133.88	28.99	79.10	107.75	136.00	153.50	180.90	98.84	1.56
4.10	AD Exp	What Winds Can Do	27	141.48	26.75	91.80	124.00	143.00	160.00	190.20	98.09	1.36
4.11*	AD Exp	Turn Waste Into Treasure	27	126.15	28.18	81.60	100.00	132.00	146.00	179.60	97.98	2.41
4.12	AD Exp	Beaks Say A Lot!	27	132.70	24.35	92.40	109.00	138.00	149.00	182.60	98.06	1.44
4.13	AD Exp	Facts of Matter	27	137.00	29.81	82.80	125.00	137.00	164.00	183.20	98.65	2.55
4.14	DN Nar	A Wild Ride for Bella	27	135.00	23.96	90.00	113.00	137.00	154.00	178.80	99.07	0.94
4.15	DN Exp	Rainbows	27	134.59	26.12	89.80	112.00	133.00	154.00	183.60	98.81	1.40
4.16	DN Exp	A New Kind of Family	27	149.78	26.81	101.20	131.00	153.00	170.00	198.40	98.69	1.06
4.M	DN	<i>DIBELS Next</i> Median	27	138.19	24.71	98.40	113.00	139.00	158.00	184.40	98.88	1.00

Fifth Grade

5.1	AD Nar	The Sky's the Limit	25	141.40	18.62	108.00	128.00	134.00	154.50	179.50	98.92	1.32
5.2*	AD Nar	The Puppet Show	25	133.40	21.11	101.00	115.50	135.00	152.00	178.60	98.75	1.75
5.3	AD Nar	A Football Fantasy	25	137.32	20.03	104.50	123.50	129.00	157.50	175.50	98.94	1.14
5.4	AD Nar	The Winner Within	25	129.72	14.55	111.00	116.50	131.00	137.00	162.00	98.61	1.18
5.5*	AD Nar	Monica in the Metro	25	125.16	15.71	102.30	109.50	123.00	136.00	152.40	96.82	2.55
5.6	AD Exp	A Woman With A Vision	25	123.84	18.45	92.60	110.50	117.00	138.50	165.80	99.38	0.88
5.7	AD Exp	What Are Sedimentary Rocks?	25	117.76	13.44	94.70	108.00	115.00	128.00	142.00	98.37	1.98

5.8	AD Exp	How Deltas Are Formed	25	140.96	18.77	112.30	126.50	138.00	156.00	175.80	98.65	1.63
5.9*	AD Exp	The Water Cycle	25	142.04	24.37	101.40	122.00	138.00	165.00	182.70	97.17	4.26
5.10*	AD Exp	The Moons in Our Solar System	25	118.88	22.24	90.30	96.00	113.00	139.50	153.80	98.70	1.17
5.11	AD Exp	The Metamorphosis of a Moth	25	149.36	19.67	121.20	131.50	149.00	167.00	181.70	98.39	1.65
5.12	AD Exp	The British Empire	25	143.76	19.46	116.00	125.00	142.00	158.00	182.70	99.14	0.94
5.13	AD Exp	An American Symbol	25	136.88	23.24	101.50	116.00	132.00	158.00	170.70	98.25	1.61
5.14	DN Exp	Build a Thermometer	25	130.36	17.71	103.60	114.00	132.00	142.00	165.70	98.75	2.42
5.15	DN Nar	How Kangaroo Got Her Pouch	25	149.80	20.53	114.50	127.50	155.00	166.50	180.90	98.14	2.47
5.16	DN Exp	An Amazing City	25	134.08	22.68	81.60	118.00	139.00	151.00	170.70	98.31	1.55
5.M	DN	<i>DIBELS Next Median</i>	25	137.76	19.43	105.40	120.00	139.00	152.00	172.50	98.11	3.24

Sixth Grade

6.1	AD Nar	Skydiving Sensation	26	128.73	18.67	102.70	113.75	127.00	142.75	165.65	97.78	2.60
6.2	AD Nar	Time to Move	26	138.04	17.73	108.05	125.50	136.00	152.00	168.60	98.18	1.84
6.3	AD Nar	Solving the Problem of Baby Face	25	130.00	20.03	103.30	113.00	126.00	141.00	173.00	98.65	1.50
6.4*	AD Nar	Butterflies in my Stomach	26	128.62	17.08	102.40	116.25	129.50	137.25	165.50	98.72	1.50
6.5*	AD Nar	Exploring Sandia Peak	26	141.12	23.14	108.35	121.50	139.50	161.00	181.95	96.72	2.49
6.6	AD Exp	What's Beneath Your Feet?	26	140.62	20.24	110.45	123.75	136.50	159.25	176.60	98.76	1.73

6.7*	AD Exp	The Difference Between an Element and a Compound	26	136.77	19.07	113.35	119.75	134.50	155.25	175.75	98.00	2.71
6.8	AD Exp	Work with a Pulley	26	139.04	22.31	99.40	118.75	139.00	155.50	178.10	99.18	1.31
6.9*	AD Exp	Fire Rocks	26	151.35	20.35	116.25	136.00	147.00	168.25	192.20	98.22	1.69
6.10	AD Exp	Mount St. Helens and the Ring of Fire	26	123.92	27.55	83.15	103.00	120.00	145.75	172.95	98.71	2.01
6.11	AD Exp	The Gulf Stream	26	132.00	20.73	101.10	115.75	133.00	151.00	173.35	99.07	1.08
6.12	AD Exp	Smart Communication?	26	138.96	24.83	100.80	116.75	136.00	156.25	183.30	97.72	1.92
6.13	AD Exp	Types of Government	26	131.88	29.19	79.80	105.75	130.00	155.50	180.50	98.98	1.44
6.14	DN Exp	Sea of Salt	26	148.65	20.92	108.05	137.00	151.50	162.25	183.20	98.79	1.56
6.15	DN Nar	Another World	26	122.88	20.91	91.70	107.00	118.50	141.50	162.60	98.67	1.75
6.16	DN Exp	The Barefoot Runner	26	135.15	21.37	96.85	122.00	132.00	154.25	175.50	97.02	2.21
6.M	DN	<i>DIBELS Next</i> Median	26	135.85	20.47	105.05	122.00	132.00	156.00	173.55	98.78	1.44

Note. N = 159. Passage type AD = *DIBELS* for Accountability Decisions proposed passage, DN = end-of-year *DIBELS Next* benchmark passage, Exp = Expository, Nar = Narrative. The *DIBELS Next* Student-Level Median is the student-level median score derived from the three *DIBELS Next* benchmark passages. For comparison purposes, the *DIBELS Next* Median is highlighted. *=Discarded passage.

Table 4

Descriptive Statistics for Daze Passages for All Grades

Passage Number	Passage Type	Title	N	Mean	SD	Percentile				
						5th	25th	50th	75th	95th
<i>Third Grade</i>										
3.1*	AD	All Aboard the Bus	26	20.60	7.75	6.70	14.50	20.00	28.25	31.80
3.2	AD	Governments at Work	26	22.96	8.31	8.55	14.00	24.25	28.13	37.65
3.3	AD	The Bluebirds	26	22.56	7.92	9.03	16.50	21.50	29.75	36.30
3.4	AD	The First Day	26	21.60	5.71	13.00	17.38	22.00	25.75	31.98
3.5*	AD	The Reason for the Seasons	26	21.38	6.04	9.88	18.38	21.25	26.25	31.65
3.6	DN	Making Chocolate	26	26.54	7.50	11.88	22.38	28.00	31.13	38.98
<i>Fourth Grade</i>										
4.1	AD	A New Friend	26	26.92	5.91	17.18	22.75	26.50	32.00	39.13
4.2	AD	Big Mack's Help	26	25.25	5.46	15.85	21.00	25.25	28.38	35.50
4.3*	AD	Helping People In Need	25	23.46	6.54	12.15	19.50	23.00	27.00	37.40
4.4	AD	Microlending	26	27.08	5.91	18.35	21.38	27.00	30.50	41.08
4.5*	AD	The Key to Map Elements	25	22.98	4.61	15.05	19.50	22.50	25.50	33.55

4.6	DN	Diamonds State Park	26	31.62	5.13	22.35	28.25	31.50	37.00	38.98
<i>Fifth Grade</i>										
5.1	AD	America's Government	25	25.92	7.33	15.75	19.50	24.50	30.00	39.70
5.2*	AD	My Canine Chorus	25	30.94	9.27	17.50	23.25	31.50	38.75	48.25
5.3*	AD	Role of British Monarchy	26	22.44	7.28	12.50	15.00	22.50	28.38	37.60
5.4	AD	Steamy Stew	27	28.69	6.66	15.90	24.50	28.50	32.00	42.10
5.5	AD	Wonders of the World	26	25.62	7.70	12.18	21.50	23.75	29.25	44.38
5.6	DN	Making Music	27	28.44	6.78	17.10	24.00	27.00	33.50	42.10
<i>Sixth Grade</i>										
6.1*	AD	A Land of Dreams	25	32.42	7.69	16.75	27.25	33.00	38.25	47.20
6.2	AD	Serving at the Shelter	24	26.60	6.71	13.38	22.63	26.25	32.38	36.75
6.3	AD	The Coral Reef	25	26.50	6.12	15.05	21.25	26.00	31.25	37.40
6.4	AD	Whirlwinds	24	27.46	6.23	15.00	23.50	27.75	31.75	38.63
6.5*	AD	Yearbook Committee	25	33.80	7.62	17.00	28.75	35.50	39.25	45.25
6.6	DN	Palo Duro Canyon State Park	25	29.86	7.39	19.50	21.25	32.50	35.75	40.95

Note. N = 104. Passage type AD = *DIBELS* for Accountability Decisions proposed passage; DN = end-of-year *DIBELS Next* benchmark passage. For fourth, fifth, and sixth grades, mean scores are inflated due to the presence of order effects. For comparison purposes, the *DIBELS Next* benchmark passage is highlighted. *=Discarded passage.

Table 5

P-Values from F-Tests for Main Effects by Measure and Grade

Grade	Passage	Passage Type (AD or DN)	Order	Content (Exp or Nar)
<i>DIBELS Oral Reading Fluency</i>				
First	.15	.44	.08	.99
Second	.64	.23	.93	.56
Third	.18	.42	.76	.06
Fourth	.03	.15	.35	.01
Fifth	.00	.98	.01	.00
Sixth	.00	.94	.07	.99
<i>Daze</i>				
Third	.06	.00	.20	--
Fourth	.00	.00	.01	--
Fifth	.00	.17	.00	--
Sixth	.00	.56	.00	--

Note. Sample size is 159 for DORF and 103 for Daze. Values reported are *p*-values from Effect *F*-tests for differences in means between the respective groups. Order was treated as a continuous variable to test for a linear trend between the order and the

score. For fifth-grade DORF and all grades for Daze, the effect tests for Passage, Passage Type, and Content were regressed on the residuals derived from a simple linear regression predicting DORF WC with order as the dependent variable.

Table 6

Correlations for the Selected First-Grade DIBELS AD Oral Reading Fluency Passages

Passage Number	Passage Type	Title	Passage Number						
			1.1	1.3	1.4	1.5	1.6	1.8	
1.1	AD Nar	Visiting Farm	--						
1.3	AD Nar	A Trip Through the Car Wash	.94	--					
1.4	AD Nar	What's in the Box?	.90	.90	--				
1.5	AD Nar	The Crows Are in the Corn	.91	.93	.94	--			
1.6	AD Nar	In the Night Sky	.93	.94	.95	.97	--		
1.8	AD Nar	A Muddy Mess	.91	.92	.94	.95	.96	--	
1.9	AD Exp	Everywhere We See Bugs	.91	.93	.91	.94	.96	.96	.90
1.10	AD Exp	Making Crepes	.92	.93	.92	.96	.96	.96	.95
1.13	AD Exp	A Frog's Life	.95	.94	.96	.94	.97	.97	.96
1.14	DN Nar	The Cocoa Stand	.91	.92	.95	.96	.94	.94	.92
1.15	DN Exp	Parts of a Tree	.88	.92	.87	.91	.91	.91	.90
1.16	DN Nar	Going to Market	.91	.91	.96	.94	.94	.94	.97

1.M	DN	--	<i>DIBELS Next</i> Median	.92	.92	.94	.95	.95	.96
				1.9	1.10	1.13	1.14	1.15	1.16
1.10	AD	Exp	Making Crepes	.96	--				
1.13	AD	Exp	A Frog's Life	.92	.94	--			
1.14	DN	Nar	The Cocoa Stand	.95	.97	.92	--		
1.15	DN	Exp	Parts of a Tree	.91	.93	.89	.91	--	
1.16	DN	Nar	Going to Market	.90	.95	.95	.94	.90	--
1.M	DN	--	<i>DIBELS Next</i> Median	.92	.97	.94	.97	.95	.97

Note. N = 26. For comparison purposes, the *DIBELS Next* Student-Level Median is highlighted. All correlations significant, $p < .001$.

Table 7

Correlations for the Selected Second-Grade DIBELS Next and AD Oral Reading Fluency Passages

Passage Number	Passage Type	Title	Passage Number						
			2.1	2.2	2.3	2.6	2.7	2.8	
2.1	AD Nar	Popsicle Stand	--						
2.2	AD Nar	Pen Pals	.89	--					
2.3	AD Nar	A New Friend	.89	.93	--				
2.6	AD Nar	Sailing	.89	.85	.84	--			
2.7	AD Nar	Lunchbox Surprise	.93	.89	.89	.88	--		
2.8	AD Nar	Grandma's Attic	.89	.83	.91	.92	.89	--	
2.9	AD Exp	Grow Your Own Salad	.94	.90	.93	.88	.90	.91	
2.11	AD Exp	Hermit Crabs and Their Shells	.93	.87	.91	.92	.89	.90	
2.12	AD Exp	Making Tacos	.94	.92	.95	.91	.91	.92	
2.14	DN Nar	Gavin's Jump	.92	.93	.91	.89	.93	.88	
2.15	DN Exp	The New Year	.92	.93	.94	.91	.91	.90	
2.16	DN Nar	Roller Skating Fun	.88	.88	.89	.89	.94	.89	

2.M	DN	--	<i>DIBELS Next</i> Median	.92	.95	.94	.91	.93	.91
				2.9	2.11	2.12	2.14	2.15	2.16
2.11	AD	Exp	Hermit Crabs and Their Shells	.93	--				
2.12	AD	Exp	Making Tacos	.95	.93	--			
2.14	DN	Nar	Gavin's Jump	.91	.91	.93	--		
2.15	DN	Exp	The New Year	.92	.95	.94	.95	--	
2.16	DN	Nar	Roller Skating Fun	.85	.89	.89	.90	.93	--
2.M	DN	--	<i>DIBELS Next</i> Median	.93	.94	.94	.97	.99	.95

Note. N = 27. For comparison purposes, the *DIBELS Next* Student-Level Median is highlighted. All correlations significant, $p < .001$.

Table 8

DIBELS AD Readability Study Correlations for the Selected Third-Grade DIBELS Next and AD Oral Reading Fluency Passages

Passage Number	Passage Type	Title	Passage Number						
			3.1	3.2	3.4	3.5	3.6	3.8	
3.1	AD Nar	Going on a Field Trip to a Nature Preserve	--						
3.2	AD Nar	A Doggie Dog World	.93	--					
3.4	AD Nar	Little Silva	.92	.92	--				
3.5	AD Nar	A New View of Vegetables	.91	.95	.90	--			
3.6	AD Nar	The Most Unusual Hats	.92	.91	.87	.92	--		
3.8	AD Nar	Pecos Bill	.91	.92	.86	.88	.82	--	
3.9	AD Exp	Kenya Has a Lot to Offer	.93	.90	.89	.89	.90	.88	
3.10	AD Exp	Africa's Natural Resources	.88	.90	.93	.90	.85	.86	
3.13	AD Exp	Henry Ford's Cars	.94	.89	.88	.88	.91	.86	
3.14	DN Nar	A Surprising Discovery	.94	.91	.95	.93	.91	.86	
3.15	DN Nar	A Day for a Shadow Dance	.90	.89	.84	.88	.87	.92	
3.16	DN Exp	A Triple Challenge	.90	.90	.86	.91	.87	.88	

3.M	DN	--	<i>DIBELS Next</i> Median	.94	.95	.93	.94	.92	.92
				3.9	3.10	3.13	3.14	3.15	3.16
3.10	AD	Exp	Africa's Natural Resources	.91	--				
3.13	AD	Exp	Henry Ford's Cars	.93	.91	--			
3.14	DN	Nar	A Surprising Discovery	.92	.93	.91	--		
3.15	DN	Nar	A Day for a Shadow Dance	.86	.77	.83	.84	--	
3.16	DN	Exp	A Triple Challenge	.93	.90	.90	.92	.84	--
3.M	DN	--	<i>DIBELS Next</i> Median	.93	.91	.91	.95	.93	.95

Note. N = 28. For comparison purposes, the *DIBELS Next* Student-Level Median is highlighted. All correlations significant, $p < .001$.

Table 9

Correlations for the Selected Fourth-Grade DIBELS Next and AD Oral Reading Fluency Passages

Passage Number	Passage Type	Title	Passage Number						
			4.1	4.3	4.4	4.5	4.6	4.9	
4.1	AD Nar	Just One Wheel?	--						
4.3	AD Nar	A History Lesson	.84	--					
4.4	AD Exp	Plants Have Needs, Too	.80	.85	--				
4.5	AD Exp	Stages of Life	.89	.82	.80	--			
4.6	AD Nar	Quite A Game	.82	.87	.84	.82	--		
4.9	AD Exp	The Bill of Rights	.81	.79	.86	.87	.82	--	
4.10	AD Exp	What Winds Can Do	.78	.89	.86	.81	.84	.79	
4.12	AD Exp	Beaks Say A Lot!	.73	.85	.87	.79	.89	.83	
4.13	AD Exp	Facts of Matter	.79	.78	.74	.80	.78	.80	
4.14	DN Nar	A Wild Ride for Bella	.87	.82	.82	.86	.82	.88	
4.15	DN Exp	Rainbows	.82	.88	.85	.88	.84	.82	
4.16	DN Exp	A New Kind of Family	.86	.84	.79	.92	.81	.80	

4.M	DN	--	<i>DIBELS Next</i> Median	.88	.89	.85	.91	.85	.83
				4.10	4.12	4.13	4.14	4.15	4.16
4.12	AD	Exp	Beaks Say A Lot!	.84	--				
4.13	AD	Exp	Facts of Matter	.78	.67	--			
4.14	DN	Nar	A Wild Ride for Bella	.82	.79	.82	--		
4.15	DN	Exp	Rainbows	.89	.89	.79	.83	--	
4.16	DN	Exp	A New Kind of Family	.88	.81	.80	.81	.89	--
4.M	DN	--	<i>DIBELS Next</i> Median	.93	.87	.82	.87	.97	.95

Note. N = 27. For comparison purposes, the *DIBELS Next* Student-Level Median is highlighted. All correlations significant, $p < .001$.

Table 10

Correlations for the Selected Fifth-Grade DIBELS Next and AD Oral Reading Fluency Passages

Passage Number	Passage Type	Title	Passage Number						
			5.1	5.3	5.4	5.6	5.7	5.8	
5.1	AD Nar	The Sky's the Limit	--						
5.3	AD Nar	A Football Fantasy	.79	--					
5.4	AD Nar	The Winner Within	.81	.85	--				
5.6	AD Exp	A Woman With A Vision	.80	.87	.88	--			
5.7	AD Exp	What Are Sedimentary Rocks?	.74	.80	.75	.76	--		
5.8	AD Exp	How Deltas Are Formed	.79	.75	.68	.59**	.70	--	
5.11	AD Exp	The Metamorphosis of a Moth	.75	.80	.75	.73	.77		.76
5.12	AD Exp	The British Empire	.78	.74	.70	.70	.68		.76
5.13	AD Exp	An American Symbol	.85	.79	.81	.83	.67		.61
5.14	DN Exp	Build a Thermometer	.78	.85	.78	.82	.80		.79
5.15	DN Nar	How Kangaroo Got Her Pouch	.73	.86	.81	.89	.70		.63
5.16	DN Exp	An Amazing City	.67	.73	.82	.81	.71		.60

5.M	DN	--	<i>DIBELS Next</i> Median	.82	.86	.88	.91	.78	.71
				5.11	5.12	5.13	5.14	5.15	5.16
5.12	AD	Exp	The British Empire	.76	--				
5.13	AD	Exp	An American Symbol	.68	.71	--			
5.14	DN	Exp	Build a Thermometer	.79	.76	.71	--		
5.15	DN	Nar	How Kangaroo Got Her Pouch	.77	.70	.78	.90	--	
5.16	DN	Exp	An Amazing City	.63	.60	.74	.70	.72	--
5.M	DN	--	<i>DIBELS Next</i> Median	.82	.74	.87	.90	.95	.84

Note. N = 25. For comparison purposes, the *DIBELS Next* Student-Level Median is highlighted. Unless marked, all correlations significant, $p < .001$; $**p < .01$.

Table 11

Correlations for the Selected Sixth-Grade DIBELS Next and AD Oral Reading Fluency Passages

Passage Number	Passage Type	Title	Passage Number						
			6.1	6.2	6.3	6.6	6.8	6.10	
6.1	AD Nar	Skydiving Sensation	--						
6.2	AD Nar	Time to Move	.77	--					
6.3	AD Nar	Solving the Problem of Baby Face	.72	.88	--				
6.6	AD Exp	What's Beneath Your Feet?	.56	.71	.76	--			
6.8	AD Exp	Works with a Pulley	.83	.80	.83	.66	--		
6.10	AD Exp	Mount St. Helens and the Ring of Fire	.82	.78	.75	.62	.80	--	
6.11	AD Exp	The Gulf Stream	.86	.77	.83	.71	.84	.84	
6.12	AD Exp	Smart Communication?	.85	.72	.70	.49**	.77	.78	
6.13	AD Exp	Types of Government	.78	.80	.83	.67	.81	.75	
6.14	DN Exp	Sea of Salt	.83	.88	.86	.65	.83	.82	
6.15	DN Exp	Another World	.84	.89	.91	.71	.87	.85	
6.16	DN Nar	The Barefoot Runner	.75	.77	.75	.54	.83	.84	

6.M	DN	--	<i>DIBELS Next</i> Median	.88	.88	.89	.68	.87	.87
				6.11	6.12	6.13	6.14	6.15	6.16
6.12	AD	Exp	Smart Communication?	.77	--				
6.13	AD	Exp	Types of Government	.88	.72	--			
6.14	DN	Exp	Sea of Salt	.86	.77	.83	--		
6.15	DN	Exp	Another World	.86	.79	.84	.86	--	
6.16	DN	Nar	The Barefoot Runner	.83	.79	.79	.76	.80	--
6.M	DN	--	<i>DIBELS Next</i> Median	.88	.84	.84	.88	.98	.85

Note. N = 26. For comparison purposes, the *DIBELS Next* Student-Level Median is highlighted. Unless marked, all correlations significant, $p < .001$; $**p < .01$.

Table 12
Correlations for Daze Passages for All Grades

Passage Number	Passage Type	Title	Correlation with Corresponding Passage		
			3.2	3.3	3.4
<i>Third Grade</i>			3.2	3.3	3.4
3.2	AD	Governments at Work	--		
3.3	AD	The Bluebirds	.78	--	
3.4	AD	The First Day	.70	.77	--
3.6	DN	Making Chocolate	.67	.75	.59
<i>Fourth Grade</i>			4.1	4.2	4.4
4.1	AD	A New Friend	--		
4.2	AD	Big Mack's Help	.53	--	
4.4	AD	Microlending	.62	.60	--
4.6	DN	Diamonds State Park	.65	.46**	.60
<i>Fifth Grade</i>			5.1	5.4	5.5
5.1	AD	America's Government	--		
5.4	AD	Steamy Stew	.74	--	
5.5	AD	Wonders of the World	.69	.70	--
5.6	DN	Making Music	.51	.59	.44*
<i>Sixth Grade</i>			6.2	6.3	6.4
6.2	AD	Serving at the Shelter	--		
6.3	AD	The Coral Reef	.58	--	
6.4	AD	Whirlwinds	.70	.59	--

6.6	DN	Palo Duro Canyon State Park	.73	.56	.65
-----	----	-----------------------------	-----	-----	-----

Note. N = 104. Sample size by grade: third grade n = 26, fourth grade n = 26, fifth grade n = 27, and sixth grade n = 25. For comparison purposes, the *DIBELS Next* benchmark passage is highlighted. Unless marked, all correlations significant, $p < .001$; ***** $p < .01$, **** $p < .05$.

Table 13

Passage-Level Residuals for the Selected DIBELS Oral Reading Fluency (DORF) Passages

Passage Number	Passage Type	Title	N	Mean	SD
<i>First Grade</i>					
1.1	Nar	Visiting Farm	26	-4.24	8.96
1.3	Nar	A Trip Through the Car Wash	26	0.95	8.27
1.4	Nar	What's in the Box?	26	6.02	8.41
1.5	Nar	The Crows Are in the Corn	26	3.19	7.33
1.6	Nar	In the Night Sky	26	1.98	4.86
1.8	Nar	A Muddy Mess	26	-6.83	7.00
1.9	Exp	Everywhere We See Bugs	26	2.34	6.21
1.10	Exp	Making Crepes	26	0.61	8.19
1.13	Exp	A Frog's Life	26	0.77	6.32
1.M	--	<i>DIBELS Next</i> Median	26	-4.44	7.18
<i>Second Grade</i>					
2.1	Nar	Popsicle Stand	27	-3.34	9.41
2.2	Nar	Pen Pals	27	3.30	11.20
2.3	Nar	A New Friend	27	-9.65	9.55
2.6	Nar	Sailing	27	7.22	9.55
2.7	Nar	Lunchbox Surprise	27	-4.02	9.82
2.8	Nar	Grandma's Attic	27	-2.43	9.94
2.9	Exp	Grow Your Own Salad	27	2.02	8.97

2.11	Exp	Hermit Crabs and Their Shells	27	-1.27	9.44
2.12	Exp	Making Tacos	27	4.45	5.61
2.M	--	<i>DIBELS Next</i> Median	27	3.84	9.15

Third Grade

3.1	Nar	Going on a Field Trip to a Nature Preserve	28	1.49	7.01
3.2	Nar	A Doggie Dog World	27	-5.65	7.17
3.4	Nar	Little Silva	28	7.81	10.73
3.5	Nar	A New View of Vegetables	28	3.61	10.12
3.6	Nar	The Most Unusual Hats	27	1.10	10.41
3.8	Nar	Pecos Bill	27	-5.02	10.02
3.9	Exp	Kenya Has a Lot to Offer	28	2.66	8.42
3.10	Exp	Africa's Natural Resources	28	-5.97	8.77
3.13	Exp	Henry Ford's Cars	28	-1.66	9.32
3.M	--	<i>DIBELS Next</i> Median	28	1.27	6.94

Fourth Grade

4.1	Nar	Just One Wheel?	27	-0.05	9.22
4.3	Nar	A History Lesson	26	1.38	10.29
4.4	Exp	Plants Have Needs, Too	27	-0.99	8.28
4.5	Exp	Stages of Life	26	-4.27	11.48
4.6	Nar	Quite A Game	27	3.23	11.05
4.9	Exp	The Bill of Rights	27	-5.83	9.32
4.10	Exp	What Winds Can Do	27	-2.64	9.13
4.12	Exp	Beaks Say A Lot!	27	2.01	11.77

4.13	Exp	Facts of Matter	27	2.75	12.73
4.M	--	<i>DIBELS Next Median</i>	27	4.30	6.66

Fifth Grade

5.1	Nar	The Sky's the Limit	25	5.75	7.48
5.3	Nar	A Football Fantasy	24	5.30	8.05
5.4	Nar	The Winner Within	26	0.16	8.95
5.6	Exp	A Woman With A Vision	25	4.29	6.95
5.7	Exp	What Are Sedimentary Rocks?	25	4.40	9.03
5.8	Exp	How Deltas Are Formed	25	0.62	9.55
5.11	Exp	The Metamorphosis of a Moth	25	-10.00	7.10
5.12	Exp	The British Empire	24	-7.89	9.37
5.13	Exp	An American Symbol	25	-2.96	8.02
5.M	--	<i>DIBELS Next Median</i>	25	0.24	6.26

Sixth Grade

6.1	Nar	Skydiving Sensation	26	-3.90	7.73
6.2	Nar	Time to Move	26	-4.60	11.89
6.3	Nar	Solving the Problem of Baby Face	26	0.96	7.33
6.6	Exp	What's Beneath Your Feet?	26	-1.70	11.46
6.8	Exp	Work with a Pulley	26	0.87	8.07
6.10	Exp	Mount St. Helens and the Ring of Fire	26	2.40	7.37
6.11	Exp	The Gulf Stream	26	4.97	8.16
6.12	Exp	Smart Communication?	25	-4.04	9.11
6.13	Exp	Types of Government	26	5.20	9.14

6.M	--	<i>DIBELS Next</i> Median	26	-0.31	5.68
-----	----	---------------------------	----	-------	------

Note. N = 159. Passage type Exp = Expository, Nar = Narrative. The *DIBELS Next* Median is the mean of the student-level median scores derived from the three *DIBELS Next* benchmark passages. For comparison purposes, the *DIBELS Next* Student-Level Median is highlighted.

Table 14

Passage-Level Residuals for the Selected Daze Passages

Passage Number	Passage Type	Title	N	Mean	SD
<i>Third Grade</i>					
3.2	AD	Governments at Work	26	-0.80	3.04
3.3	AD	The Bluebirds	26	-0.29	2.61
3.4	AD	The First Day	26	-0.59	2.34
3.6	DN	Making Chocolate	26	1.67	3.27
<i>Fourth Grade</i>					
4.1	AD	A New Friend	26	-0.03	2.11
4.2	AD	Big Mack's Help	26	-1.38	2.55
4.4	AD	Microlending	26	-0.55	3.14
4.6	DN	Diamonds State Park	26	1.96	3.09
<i>Fifth Grade</i>					
5.1	AD	America's Government	25	-0.93	2.20
5.4	AD	Steamy Stew	27	0.95	2.88
5.5	AD	Wonders of the World	26	-1.59	2.67
5.6	DN	Making Music	27	1.44	3.39
<i>Sixth Grade</i>					
6.2	AD	Serving at the Shelter	24	-0.24	2.36
6.3	AD	The Coral Reef	25	-0.36	2.47
6.4	AD	Whirlwinds	24	-0.73	2.42

6.6	DN	Palo Duro Canyon State Park	25	1.29	2.95
-----	----	-----------------------------	----	------	------

Note. N = 104. Passage type AD = *DIBELS* for Accountability Decisions proposed passage; DN = end-of-year *DIBELS Next* benchmark passage. For comparison purposes, the *DIBELS Next* benchmark passage is highlighted.

Table 15

Descriptive Statistics for the Selected DORF Triads and Paired Daze Passage

Triad by Grade	DIBELS Oral Reading Fluency (DORF)						Paired Daze Passage		
	Passage Number			Mean	SD	Σres	Passage Number	Mean	SD
	1st	2nd	3rd						
<i>First Grade</i>									
Triad 1	1.3	1.6	1.10	61.69	32.34	1.95	--	--	--
Triad 2	1.1	1.5	1.9	60.94	31.12	0.30	--	--	--
Triad 3	1.4	1.8	1.13	61.13	31.63	1.63	--	--	--
DN Triad	--	--	--	57.15	28.80	--	--	--	--
<i>Second Grade</i>									
Triad 1	2.3	2.6	2.9	100.29	34.17	0.82	--	--	--
Triad 2	2.1	2.2	2.11	99.01	34.65	0.08	--	--	--
Triad 3	2.7	2.8	2.12	97.88	32.65	0.77	--	--	--
DN Triad	--	--	--	103.38	38.21	--	--	--	--
<i>Third Grade</i>									
Triad 1	3.2	3.4	3.13	119.02	30.54	1.04	3.2	22.96	8.31
Triad 2	3.1	3.5	3.10	119.10	32.36	0.33	3.4	21.60	5.71
Triad 3	3.6	3.8	3.9	118.22	30.83	0.72	3.3	22.56	7.92
DN Triad	--	--	--	121.05	26.54	--	3.6	26.54	7.50
<i>Fourth Grade</i>									
Triad 1	4.4	4.5	4.6	132.26	26.74	0.56	4.4	27.08	5.91

Triad 2	4.1	4.10	4.12	132.40	26.33	0.79	4.2	25.25	5.46
Triad 3	4.3	4.9	4.13	132.13	27.10	0.23	4.1	26.92	5.91
DN Triad	--	--	--	139.79	25.66	--	4.6	26.54	5.13

Fifth Grade

Triad 1	5.1	5.8	5.12	136.58	17.87	1.4	5.4	28.69	6.66
Triad 2	5.3	5.6	5.11	135.90	19.93	0.29	5.1	25.92	7.33
Triad 3	5.4	5.7	5.13	137.43	21.17	1.72	5.5	25.62	7.70
DN Triad	--	--	--	137.76	19.43	--	5.6	28.44	6.78

Sixth Grade

Triad 1	6.3	6.8	6.11	135.67	21.07	0.74	6.4	27.46	6.71
Triad 2	6.2	6.7	6.13	137.18	24.22	1.41	6.3	26.50	6.12
Triad 3	6.1	6.6	6.10	136.63	21.05	0.69	6.2	26.60	6.23
DN Triad	--	--	--	135.85	20.47	--	6.6	29.86	7.39

Note. N = 159. $|\Sigma\text{res}|$ is the absolute value of the sum of the passage-level residuals for that triad. DN Triad = *DIBELS Next* end-of-year benchmark triad.

Table 16

DORF Triad Correlations

Grade	Correlations Between Triads			Correlations by Triad with the <i>DIBELS</i> Next End-of-Year Benchmark Triad		
	1 - 2	1 - 3	2 - 3	1	2	3
First	.99	.97	.97	.98	.97	.97
Second	.97	.98	.96	.96	.97	.96
Third	.98	.97	.97	.96	.97	.98
Fourth	.97	.95	.94	.95	.97	.94
Fifth	.85	.85	.93	.82	.93	.91
Sixth	.91	.93	.90	.96	.94	.92

Note. N = 159. All correlations significant, $p < .001$.

Table 17

Comparison of the DORF Triads to the DIBELS Next End-of-Year Benchmark Triad

Triad by Grade	Mean	SD	$ \Sigma res $	P-value for Difference from DNext Triad
<i>First Grade</i>				
Triad 1	61.69	32.34	1.95	.00
Triad 2	60.94	31.12	0.30	.03
Triad 3	61.13	31.63	1.63	.06
DN Triad	57.15	28.80	--	--
<i>Second Grade</i>				
Triad 1	100.29	34.17	0.82	.21
Triad 2	99.01	34.65	0.08	.06
Triad 3	97.88	32.65	0.77	.05
DN Triad	103.38	38.21	--	--
<i>Third Grade</i>				
Triad 1	119.02	30.54	1.04	.33
Triad 2	119.10	32.36	0.33	.21
Triad 3	118.22	30.83	0.72	.14
DN Triad	121.05	26.54	--	--
<i>Fourth Grade</i>				
Triad 1	132.26	26.74	0.56	.00
Triad 2	132.40	26.33	0.79	.00
Triad 3	132.13	27.10	0.23	.00

DN Triad	139.79	25.66	--	--
<i>Fifth Grade</i>				
Triad 1	136.58	17.87	1.4	.60
Triad 2	135.90	19.93	0.29	.19
Triad 3	137.43	21.17	1.72	.55
DN Triad	137.76	19.43	--	--
<i>Sixth Grade</i>				
Triad 1	135.67	21.07	0.74	.00
Triad 2	137.18	24.22	1.41	.20
Triad 3	136.63	21.05	0.69	.98
DN Triad	135.85	20.47	--	--

Note. N = 159. $|\Sigma\text{res}|$ is the absolute value of the sum of the passage-level residuals for that triad. *P*-value calculated from a matched-pairs test for a difference from the *DIBELS Next* end-of-year benchmark triad. DN Triad = *DIBELS Next* end-of-year benchmark triad.

Table 18

Comparison of the Daze Passage to the DIBELS Next End-of-Year Benchmark Daze Passage

Grade	N	Passage	Mean Difference	95% CI for the Difference	P-value for Difference from DNext Passage
Third	26	3.2	3.58	(0.97, 6.18)	0.01
Fourth	25	4.4	5.09	(3.25, 6.90)	0.00
Fifth	27	5.4	-0.1	(-2.4, 1.13)	0.94
Sixth	24	6.4	2.98	(1.24, 5.53)	0.02

Note. P-values reported from a matched-pairs test for differences in Daze Adjusted Score. For fourth, fifth, and sixth grade, the Daze Adjusted Score was corrected for effects of passage order.

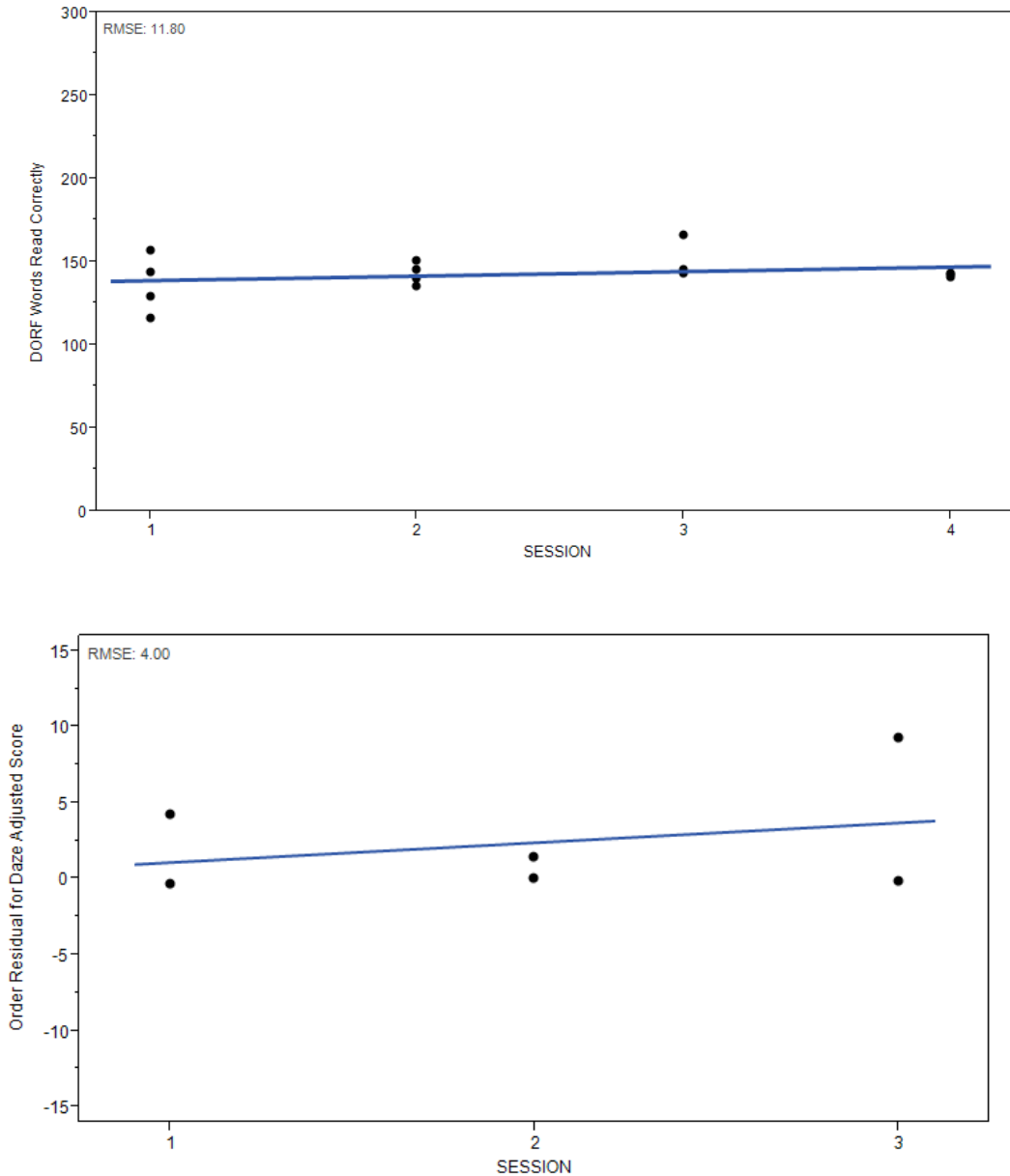


Figure 1. Example scatter plots. The top plot is from one fourth-grade student’s DIBELS Oral Reading Fluency (DORF) passage scores (marked as large dots) plotted across four testing sessions with a simple linear regression line. The bottom plot is from the same fourth-grade student’s Daze passage scores (adjusted for order effects) plotted across the three testing sessions with a simple linear regression line. For both measures, the passage-level residual is the distance between each individual passage score and the line at each testing session.

Appendix

Table A1

First-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD DORF Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
Visiting a Farm	1.1	Nar	218.00	7.5	-1.51	-0.75	-0.97	-1.08
Catching Frogs*	1.2	Nar	248.00	9	-1.02	-1.30	-1.41	-1.24
A Trip Through the Car Wash	1.3	Nar	248.00	8.5	-1.18	-1.15	-0.94	-1.09
What's in the Box?	1.4	Nar	249.00	8	-1.35	-1.55	-1.16	-1.35
The Crows Are in the Corn	1.5	Nar	249.00	9.5	-0.86	-1.62	-1.25	-1.24
In the Night Sky	1.6	Nar	244.00	7.5	-1.51	-0.88	-0.86	-1.08
A Thrilling Ride*	1.7	Nar	237.00	8	-1.35	-0.83	-1.25	-1.14
A Muddy Mess	1.8	Nar	201.00	8	-1.35	-0.78	-1.63	-1.25
Everywhere We See Bugs	1.9	Exp	230.00	8	-1.35	-1.11	-0.99	-1.15
Making Crepes	1.10	Exp	222.00	8.5	-1.18	-1.49	-1.42	-1.36
Wind and Rain*	1.11	Exp	209.00	9	-1.02	-1.13	-1.26	-1.14
Homes, Homes, Homes*	1.12	Exp	250.00	8	-1.35	-1.39	-1.33	-1.36

A Frog's Life	1.13	Exp	209.00	7.5	-1.51	-0.76	-0.97	-1.08
Mean			231.85	8.23	-1.27	-1.13	-1.19	-1.20
Standard Deviation			18.03	0.63	0.21	0.32	0.23	0.11

*=Discarded Passages

Table A2

Second-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD DORF Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
Popsicle Stand	2.1	Nar	276.00	9.00	-1.02	-0.67	-0.38	-0.69
Pen Pals	2.2	Nar	263.00	9.00	-1.02	-0.71	-1.00	-0.91
A New Friend	2.3	Nar	275.00	9.00	-1.02	-0.73	-0.9	-0.88
Firefly Night*	2.4	Nar	247.00	9.00	-1.02	-0.35	-0.34	-0.57
Bakery Bear*	2.5	Nar	274.00	9.00	-1.02	-0.80	-0.94	-0.92
Sailing	2.6	Nar	273.00	10.00	-0.69	-0.50	-0.63	-0.61
Lunchbox Surprise	2.7	Nar	268.00	9.00	-1.02	-0.75	-0.33	-0.70
Grandma's Attic	2.8	Nar	274.00	10.50	-0.53	-1.08	-0.87	-0.83
Grow Your Own Salad	2.9	Exp	275.00	9.00	-1.02	-0.73	-0.79	-0.85
What Causes the Seasons?*	2.10	Exp	270.00	9.50	-0.86	-0.91	-0.71	-0.83
Hermit Crabs and Their Shells	2.11	Exp	265.00	10.50	-0.53	-0.73	-0.98	-0.75

Making Tacos	2.12	Exp	238.00	9.00	-1.02	-0.34	-0.52	-0.63
Splish, Splash!*	2.13	Exp	270.00	10.00	-0.69	-0.91	-0.38	-0.66
Mean			266.77	9.42	-0.88	-0.71	-0.67	-0.76
Standard Deviation			11.63	0.61	0.20	0.21	0.26	0.12

*=Discarded Passages

Table A3

Third-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD DORF Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
Going on a Field Trip to a Nature Preserve	3.1	Nar	269.00	12.00	-0.04	-0.69	-0.56	-0.43
A Doggie Dog World	3.2	Nar	301.00	11.50	-0.20	0.27	-0.65	-0.19
Tool Shed Surprise*	3.3	Nar	300.00	11.50	-0.20	0.21	-0.28	-0.09
Little Silva	3.4	Nar	300.00	11.00	-0.37	-0.31	-0.28	-0.32
A New View of Vegetables	3.5	Nar	295.00	12.00	-0.04	-0.07	-0.18	-0.10
The Most Unusual Hats	3.6	Nar	267.00	11.00	-0.37	0.06	0.05	-0.09
Icing on the Cake*	3.7	Nar	300.00	11.00	-0.37	0.08	-0.01	-0.10
Pecos Bill	3.8	Nar	264.00	11.00	-0.37	0.02	-0.60	-0.32
Kenya Has a Lot to Offer	3.9	Exp	269.00	11.50	-0.20	0.26	-0.08	-0.01
Africa's Natural Resources	3.10	Exp	254.00	11.00	-0.37	-0.18	0.28	-0.09
Geography of Crete*	3.11	Exp	251.00	11.00	-0.37	0.24	0.05	-0.03

Sink or Float?*	3.12	Exp	300.00	11.50	-0.20	-0.05	-0.28	-0.18
Henry Ford's Cars	3.13	Exp	300.00	12.00	-0.04	0.15	-0.13	-0.01
Mean			282.31	11.38	-0.24	-0.001	-0.21	-0.15
Standard Deviation			19.96	0.42	0.14	0.27	0.28	0.13

*=Discarded Passages

Table A4

Third-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD Daze Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
All Aboard the Bus*	3.1	Nar	395.00	12.00	-0.04	-0.15	-0.25	-0.15
Governments at Work	3.2	Exp	355.00	11.00	-0.37	-0.08	0.19	-0.09
The Bluebirds	3.3	Nar	392.00	11.00	-0.37	-0.13	-0.11	-0.22
The First Day	3.4	Nar	399.00	11.00	-0.37	-0.17	-0.60	-0.38
The Reason for the Seasons*	3.5	Exp	391.00	11.00	-0.37	-0.67	0.27	-0.26
Mean			386.40	11.20	-0.30	-0.24	-0.10	-0.22
Standard Deviation			17.83	0.45	0.15	0.24	0.35	0.11

*=Discarded Passages

Table A5

Fourth-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD DORF Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
Just One Wheel?	4.1	Nar	349.00	12.50	0.13	0.36	0.42	0.30
Back to Earth*	4.2	Nar	328.00	13.00	0.29	0.56	0.06	0.30
A History Lesson	4.3	Nar	332.00	12.50	0.13	-0.13	0.80	0.27
Plants Have Needs, Too	4.4	Exp	346.00	13.00	0.29	-0.18	0.76	0.29
Stages of Life	4.5	Exp	343.00	14.00	0.62	0.19	0.55	0.45
Quite A Game	4.6	Nar	339.00	12.00	-0.04	0.05	0.50	0.17
Cat in the Box*	4.7	Nar	347.00	13.00	0.29	0.32	0.21	0.27
Telescopes*	4.8	Exp	340.00	14.00	0.62	0.27	0.41	0.43
The Bill of Rights	4.9	Exp	350.00	13.00	0.29	0.35	0.49	0.38
What Winds Can Do	4.10	Exp	333.00	12.50	0.13	0.04	0.15	0.11
Turn Waste Into Treasure*	4.11	Exp	350.00	12.00	-0.04	0.46	0.55	0.32
Beaks Say A Lot!	4.12	Exp	342.00	13.00	0.29	0.59	0.21	0.36

Facts of Matter	4.13	Exp	346.00	13.00	0.29	-0.07	0.28	0.17
Mean			337.75	13.03	0.30	0.19	0.34	0.28
Standard Deviation			14.04	0.67	0.22	0.24	0.27	0.10

*=Discarded Passages

Table A6

Fourth-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD Daze Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
A New Friend	4.1	Nar	449.00	13.00	0.29	0.15	0.16	0.20
Big Mack's Help	4.2	Nar	431.00	12.00	-0.04	0.32	0.45	0.24
Helping People In Need*	4.3	Exp	399.00	12.00	-0.04	0.32	0.80	0.36
Microlending	4.4	Exp	435.00	14.00	0.62	-0.07	0.80	0.45
The Key to Map Elements*	4.5	Exp	434.00	14.00	0.62	0.39	0.27	0.43
Mean			430.50	13.00	0.29	0.25	0.42	0.32
Standard Deviation			16.66	0.89	0.30	0.18	0.33	0.11

*=Discarded Passages

Table A7

Fifth-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD DORF Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
The Sky's the Limit	5.1	Nar	349.00	13.00	0.29	0.86	0.68	0.61
The Puppet Show*	5.2	Nar	348.00	14.00	0.62	0.65	0.52	0.60
A Football Fantasy	5.3	Nar	350.00	13.00	0.29	0.85	0.53	0.56
The Winner Within	5.4	Nar	349.00	15.00	0.95	0.30	1.00	0.75
Monica in the Metro*	5.5	Nar	344.00	15.00	0.95	0.75	0.68	0.79
A Woman With A Vision	5.6	Exp	331.00	13.00	0.29	0.82	1.09	0.73
What Are Sedimentary Rocks?	5.7	Exp	347.00	14.00	0.62	0.66	0.56	0.61
How Deltas Are Formed	5.8	Exp	350.00	13.50	0.45	1.08	0.23	0.59
The Water Cycle*	5.9	Exp	345.00	15.00	0.95	0.22	0.93	0.70
The Moons in Our Solar System*	5.10	Exp	339.00	14.00	0.62	1.09	0.80	0.84
The Metamorphosis of a Moth	5.11	Exp	321.00	15.00	0.95	0.63	0.72	0.77

The British Empire	5.12	Exp	338.00	13.00	0.29	1.04	0.82	0.72
An American Symbol	5.13	Exp	350.00	14.00	0.62	1.19	0.72	0.84
Mean			343.15	13.96	0.61	0.78	0.71	0.70
Standard Deviation			8.82	0.83	0.27	0.29	0.23	0.10

*=Discarded Passages

Table A8

Fifth-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD Daze Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
America's Government	5.1	Exp	480.00	14.00	0.62	0.50	1.18	0.77
My Canine Chorus*	5.2	Nar	490.00	15.00	0.95	0.35	0.65	0.65
Role of British Monarchy*	5.3	Exp	476.00	13.00	0.29	1.02	1.12	0.81
Steamy Stew	5.4	Nar	498.00	14.50	0.78	1.01	0.73	0.84
Wonders of the World	5.5	Exp	495.00	14.50	0.78	0.91	0.84	0.84
Mean			489.83	14.00	0.62	0.82	0.93	0.79
Standard Deviation			9.85	0.84	0.27	0.32	0.22	0.07

*=Discarded Passages

Table A9

Sixth-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD DORF Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
Skydiving Sensation	6.1	Nar	349.00	16.50	1.44	1.54	0.76	1.25
Time to Move	6.2	Nar	340.00	16.00	1.27	1.30	0.99	1.19
Solving the Problem of Baby Face	6.3	Nar	339.00	16.00	1.27	1.60	0.64	1.17
Butterflies in my Stomach*	6.4	Nar	346.00	16.00	1.27	0.73	0.84	0.95
Exploring Sandia Peak*	6.5	Nar	350.00	15.00	0.95	1.30	0.70	0.98
What's Beneath Your Feet?	6.6	Exp	302.00	16.00	1.27	1.36	1.04	1.22
The Difference Between an Element and a Compound*	6.7	Exp	337.00	15.00	0.95	0.76	1.57	1.09
Work with a Pulley	6.8	Exp	338.00	15.00	0.95	0.98	0.90	0.94
Fire Rocks*	6.9	Exp	344.00	15.00	0.95	0.80	1.60	1.12
Mount St. Helens and the Ring of Fire	6.10	Exp	350.00	16.00	1.27	1.52	0.89	1.23
The Gulf Stream	6.11	Exp	327.00	15.50	1.11	1.40	1.15	1.22

Smart Communication?	6.12	Exp	349.00	15.00	0.95	1.31	1.42	1.23
Types of Government	6.13	Exp	310.00	15.00	0.95	0.68	1.59	1.07
Mean			337.00	15.54	1.12	1.18	1.08	1.13
Standard Deviation			15.34	0.56	0.18	0.34	0.35	0.11

*=Discarded Passages

Table A10

Sixth-Grade Passage Difficulty Variables and DMG Passage Difficulty Index for DIBELS AD Daze Passages

Title	Passage Number	Type	Number of Words	Median Words per Sentence	Median Words per Sentence Indicator	Rare Words Indicator	Word Length Indicator	DMG Passage Difficulty Index Indicator
A Land of Dreams*	6.1	Exp	516.00	14.50	0.78	0.61	1.59	0.99
Serving at the Shelter	6.2	Nar	541.00	15.50	1.11	0.89	1.23	1.08
The Coral Reef	6.3	Exp	513.00	15.00	0.95	1.55	1.25	1.25
Whirlwinds	6.4	Exp	545.00	16.00	1.27	0.69	0.86	0.94
Yearbook Committee*	6.5	Nar	544.00	16.00	1.27	1.05	0.91	1.08
Mean			531.80	15.40	1.08	0.96	1.17	1.07
Standard Deviation			15.90	0.65	0.21	0.37	0.30	0.12

*=Discarded Passages