

2012-2013 DIBELS Data System Update *Part I: DIBELS Next Composite Score*

DIBELS Next was released for school use and supported on the DIBELS Data System (DDS) in 2010. Over the course of the past two years, researchers at the University of Oregon, Center on Teaching and Learning (UO-CTL) have been actively collecting feedback from schools about their use of the DIBELS measures with their students. Additionally, we have been conducting a series of rigorous, national evaluations focused on the technical features of DIBELS Next. The results of our research, combined with feedback from users, have prompted us to make some important changes to the DDS and the services that we offer *beginning in August, 2012*. This Technical Brief is the first in a series of documents that we are creating to help explain the rationale and research basis for the upcoming system advances.

The first important user update is that the required administration of all DIBELS Next measures (including Retell Fluency) during benchmark assessment will now be *optional* in the DDS. Of course, you will be able to continue to administer all DIBELS measures if you wish. But, you will also have the choice to administer a subset of the standard DIBELS measures with no loss of functionality in your reports and no loss of value in predicting your students' future reading performance.

We outline the recommended individual DIBELS measures for benchmarking assessment in Table 3 of this Technical Brief. Consistent with nearly forty years of educational research, our revised recommendations *always* include the administration of Oral Reading Fluency (ORF). For students in grades 3 – 6, we have concluded that administering the Daze measure should be *optional*, but we endorse its use. For students in grades 1 – 6 we have concluded that administering the Retell Fluency measure should be *optional*, and we do *not* endorse its use. We acknowledge that student skills in retelling what they have read is an important component of reading comprehension (e.g., University of Oregon Big Ideas in Beginning Reading, *Curriculum Maps for Grade 3, Focus 4: Retelling, Summarizing, Synthesizing*: http://reading.uoregon.edu/resources/maps.php#3_comp) but, at this point, the Retell Fluency measure simply does not have the technical features necessary to require its use, large-scale, for benchmark assessment.

We believe there are a number of advantages in administering a core and streamlined set of DIBELS measures for universal screening assessment (i.e., benchmarking). Perhaps the most important advantage is the amount of time schools will save on test administration. In an average K-6 school (NCES, 2011), with 52 students per grade, omitting the Retell Fluency measure will save more than 5 hours of testing per year in grade 1 and approximately 8 hours of testing per year in grades 2 – 6. Omitting the Retell Fluency measure alone would *save an average-sized school more than 44 hours of testing time per year that could potentially be used for instruction*. And this cost-savings is stated just in terms of raw assessment administration time. Additional savings would occur in terms of reduced assessor time spent in training on Retell administration, less data entry, and more focused progress monitoring assessments for struggling readers. We believe these potential instructional savings are critical, because the use of Retell Fluency is simply not technically defensible.

What does this change mean for the Composite Score?

The DDS started supporting research on the DIBELS Next measures in 2007, allowing data collection on new measures with so-called "beta test" participants. The DDS implemented wide-scale reporting for DIBELS Next in the fall of 2010 to allow *all* DDS customers an opportunity to try out the new

measures. We thought this step was important because DIBELS Next represents the 7th edition of DIBELS. Our goal in supporting DIBELS 7th was to support continued advances in the DIBELS measures, which has been a priority for us since the DDS was first introduced to the public in 2000. Additionally, as improvements are made to DIBELS, we want you as our customers to have access to those changes.

The use of the Composite Score in the 7th edition was an important change and we set out to study it carefully. We have conducted a comprehensive study to determine if the scientific evidence supports the decision that the Composite Score is *required* to obtain a reliable predictor of students' reading proficiency. Our primary conclusion is that the scientific evidence does *not* support the requirement that all DIBELS Next measures be administered.

We base this conclusion on a national study we conducted on DIBELS Next from 2010 through 2012—subsequent to the initial DIBELS Next release in May 2010. It is important to note that *the University of Oregon did not* conduct the preliminary study used to derive initial evidence for DIBELS Next. That initial study was relatively small, and the participating schools were not representative of DDS users, or even of U.S. schools as a whole. In that preliminary study, the average number of students tested per grade (K–6) was 180, 91% of the sample was White, and only 16% of students in the sample qualified for free or reduced-price lunch. In the follow-up, UO-CTL study, we partnered with 28 schools from across 15 U.S. states to test an average of 543 students per grade. In this formal study, the demographic make-up of students in the sample closely matched all schools in the DDS and across the nation as a whole. In Table 1, we list the demographic composition of students in (a) the *preliminary* study sample (participation year 2009-2010), (b) the UO-CTL study sample (2010-2011), (c) DDS schools (2010-2011), and (d) students in the nation as a whole (2010-2011).

Table 1. *Demographic characteristics of student participants in DIBELS Next research samples*

| Demographic Category | % in Preliminary Sample | % in UO-CTL DIBELS Next Test Sample | % in All Schools in DDS | % in All Schools in the Nation |
|----------------------------------|-------------------------|-------------------------------------|-------------------------|--------------------------------|
| Female | 50.0 | 48.0 | 47.9 | 47.7 |
| American Indian / Alaskan Native | 1.0 | 4.0 | 2.9 | 2.2 |
| Asian / Pacific Islander | <1.0 | 1.1 | 3.2 | 4.3 |
| Hispanic | 6.0 | 19.7 | 14.8 | 20.0 |
| Black | <1.0 | 14.4 | 14.5 | 16.5 |
| White | 91.0 | 50.9 | 63.8 | 55.1 |
| Free / Reduced-Price Lunch | 16.0 | 63.7 | 53.3 | 52.4 |

Note. Values reported are *median* values for each sample across grades K – 6 and may not sum to 100%. Free/Reduced-Price lunch percentages for the *Preliminary Sample* are reported at the District level.

Data analysis. We used the following procedures to analyze the potential value-added of Composite Score. All participating schools administered the DIBELS Next measures according to standardized procedures and timelines for benchmark testing during the 2010-2011 academic year. All schools submitted their winter score sheets to the University of Oregon for re-scoring by UO-CTL staff in the summer of 2011. The inter-scorer agreement for all DIBELS winter measures, K – 6, was 99%.

In order to examine the *value* that a measure like Retell adds to our predictions about student reading skills, we needed to identify a critically important outcome to predict at the end of the school year. We selected the *Stanford Achievement Test—10th Edition* (SAT10; Harcourt Assessment, 2004, 2007 Normative Update) as the outcome measure that would indicate "healthy" reading performance at the end of the school year. The SAT10 is a group-administered achievement test that samples skill development in several different academic areas. For our sample, we used the Stanford *Early School Achievement Test* (SESAT) used to assess students in kindergarten, and the SAT10 to assess students in grades 1–6. Selected reading subtests from the SAT10 were combined to form a total reading composite, and it was this Total Reading Score that served as the standard for healthy reading performance in the spring of each grade.

Participating schools administered the SAT10, and UO-CTL staff trained all test administrators using a standardized protocol. We checked the fidelity of SAT10 administration with a 10-point administration checklist. Fidelity was between 92% and 99% across all test administrations, and all test protocols were considered valid for analysis.

Results. Our first question was focused on determining the *percent of value* that the DIBELS Next Composite Score added to our predictions about student SAT10 performance. We used what is called a *sequential regression procedure* to obtain our results. The generalized findings from that study are presented in Table 2.

The first two columns of numbers in Table 2 represent correlations between the single, recommended DIBELS measure with the SAT10, and then the correlations between the DIBELS Next Composite Score with the SAT10. For example, in the beginning of second grade, the correlation between ORF and the SAT10 is .680 and the correlation between the DIBELS Composite and the SAT10 is .695.

The final column in Table 2 represents the percent of additional variance in the SAT10 that is explained when *all* of the DIBELS Next measures are administered in order to obtain the Composite Score. Looking at second grade, we see that the Composite Score at the beginning of the year adds approximately 2% compared to ORF alone when predicting students' end of year SAT10 performance. At the end of the year, however, the Composite Score actually predicts the SAT10 less well than ORF by itself (i.e., ORF explains 4.31% *more* in SAT10 performance than the Composite Score).

Table 2. *Comparison of the primary DIBELS Next measures and the DIBELS Next Composite Score as they predict end-of-grade SAT10 outcomes*

| Grade | DIBELS Individual Measure Score (see note) Predicting SAT10 Total Reading | DIBELS Composite Score Predicting SAT10 Total Reading | Additional Variance Explained by DIBELS Composite Score |
|--------------------------------|---|---|---|
| Kindergarten | | | |
| Beginning of Year ^a | .592 | .639 | 5.79% |
| End of Year ^b | .647 | .671 | 3.16% |
| First Grade | | | |
| Beginning of Year ^b | .582 | .586 | 0.47% |
| End of Year ^c | .787 | .783 | –0.63% |
| Second Grade | | | |
| Beginning of Year ^c | .680 | .695 | 2.06% |
| End of Year ^c | .734 | .704 | –4.31% |

| Grade | DIBELS Individual Measure Score ^(see note) Predicting SAT10 Total Reading | DIBELS Composite Score Predicting SAT10 Total Reading | Additional Variance Explained by DIBELS Composite Score |
|--------------------------------|--|---|---|
| Third Grade | | | |
| Beginning of Year ^c | .674 | .706 | 4.42% |
| End of Year ^c | .687 | .685 | -0.27% |
| Fourth Grade | | | |
| Beginning of Year ^c | .612 | .634 | 2.74% |
| End of Year ^c | .603 | .579 | -2.84% |
| Fifth Grade | | | |
| Beginning of Year ^c | .675 | .705 | 4.14% |
| End of Year ^c | .627 | .622 | -0.62% |
| Sixth Grade | | | |
| Beginning of Year ^c | .636 | .656 | 2.58% |
| End of Year ^c | .619 | .648 | 3.67% |

Note. The superscripts specified for the Beginning of Year (BOY) and End of Year (EOY) for each grade level designate the following *individual* DIBELS Next measure recommended for primary interpretation for a given time period (in terms of its SAT10 prediction)

^a = Letter Naming Fluency;

^b = Nonsense Word Fluency—Correct Letter Sounds;

^c = Oral Reading Fluency—Words Read Correct.

Overall, we see that ORF alone is a very strong predictor in Grades 1 – 6. When administered in the *fall* (for grades 2 – 6), ORF explains approximately 40% of the variance in your students' *spring* SAT10 Total Reading Score. Given the complexity of the reading tasks that are associated with a test as comprehensive as the SAT10, we feel this predictive power is unmatched when considering a brief, reliable measure such as ORF for your screening purposes.

It has been difficult for us to find a measure that adds much to ORF's prediction—even when all of the DIBELS measures are aggregated to form a composite. Given that obtaining the Composite Score represents a considerable increase in the number of measures administered and the corresponding total time required for test administration, we find that our results make it difficult to justify recommending that you administer additional screening measures for benchmarking.

You should feel confident that the power of your prediction when screening your students is valid and reliable when using ORF alone.

Recommended measures for administration. Based on our complete analysis, we provide a set of recommendations that specify which DIBELS measures should be: (a) required for benchmark testing administration, (b) optional, but with our endorsement, and (c) optional, *without* our endorsement (see Table 3).

Table 3. *Recommended and optional DIBELS measures in the updated (2012) framework*

| Grade | Required | Optional, Endorsed | Optional, Not Endorsed |
|---------------------|---------------------------------------|--------------------|------------------------|
| Kindergarten | | | |
| Beginning of Year | FSF LNF | | |
| Middle of Year | FSF LNF PSF NWF (CLS) | NWF (WWR) | |
| End of Year | LNF NWF (CLS & WWR) | PSF | |
| First Grade | | | |
| Beginning of Year | LNF NWF (CLS & WWR) | | PSF |
| Middle of Year | ORF (WRC & Errors) NWF (CLS & WWR) | | RTF |
| End of Year | ORF (WRC & Errors) NWF (CLS & WWR) | | RTF |
| Second Grade | | | |
| Beginning of Year | ORF (WRC & Errors) NWF (CLS & WWR) | | RTF |
| Middle of Year | ORF (WRC & Errors) | | RTF |
| End of Year | ORF (WRC & Errors) | | RTF |
| Third Grade | | | |
| Beginning of Year | ORF (WRC & Errors) | Daze | RTF |
| Middle of Year | ORF (WRC & Errors) | Daze | RTF |
| End of Year | ORF (WRC & Errors) | Daze | RTF |
| Fourth Grade | | | |
| Beginning of Year | ORF (WRC & Errors) | Daze | RTF |
| Middle of Year | ORF (WRC & Errors) | Daze | RTF |
| End of Year | ORF (WRC & Errors) | Daze | RTF |
| Fifth Grade | | | |
| Beginning of Year | ORF (WRC & Errors) | Daze | RTF |
| Middle of Year | ORF (WRC & Errors) | Daze | RTF |
| End of Year | ORF (WRC & Errors) | Daze | RTF |
| Sixth Grade | | | |
| Beginning of Year | ORF (WRC & Errors) | Daze | RTF |
| Middle of Year | ORF (WRC & Errors) | Daze | RTF |
| End of Year | ORF (WRC & Errors) | Daze | RTF |

Note.

FSF = First Sound Fluency

PSF = Phonemic Segmentation Fluency

CLS = Correct Letter Sounds

ORF = Oral Reading Fluency

RTF = Retell Fluency

LNF = Letter Naming Fluency

NWF = Nonsense Word Fluency

WWR = Whole Words Read

WRC = Words Read Correct

Summary

DIBELS was created at the University of Oregon largely through federal and public grant dollars. The Center on Teaching and Learning is a UO research and outreach unit that has managed the DDS for 14 years. To this point, we have focused our efforts on building reports for DIBELS measures and we built reports for the 7th edition of DIBELS (DIBELS Next) in the same way we did for the previous versions of DIBELS.

We have determined that we can provide better services to our customers, and to the field, if we expand our efforts. Our decision to expand our efforts is based on our commitment to public service and outreach. We believe firmly that public schools own a stake in DIBELS through their participation in federally funded research. We believe that we must continue to offer not-for-profit interpretation for your DIBELS data, as well as continue to safeguard your student information in a way that is only possible through federal assurances. Toward that end, we see many exciting opportunities on the horizon, coupled with our concerns regarding current practice. For example, we have been very concerned that a number of *for-profit companies* have chosen to provide inaccurate and misleading information about the availability of DIBELS 6th Edition, stating it would no longer be available after a particular date. This statement is simply not true, it has never been true, and the University of Oregon is doing its best to correct the record.

We have also remained concerned about the Composite Score requirement for DIBELS Next interpretation. This requirement represents a significant departure from previous practice in streamlined screening assessment with Curriculum-Based Measures (CBMs). At a minimum, we feel it necessary to allow the scientific evidence to direct our response to this mandate. We think the evidence is clear that the Composite Score should *not* be required, and we will make the composite optional for obtaining DIBELS Next reports from the DDS.

Given the UO-CTL wants to play a more expanded role in DIBELS, we are consolidating our plans to move in two major directions. First, we will expand our “advisory” role on the use of DIBELS measures. UO-CTL will make decisions about which measures to give, how to improve the measures, and how to best interpret student performance. You have our assurance that we will make all of these decisions based on sound scientific methods and evidence. Second, in collaboration with selected partners, we will provide expanded services in the proper administration and scoring of DIBELS and how you can best use DIBELS measures and data once it’s collected.

We will have much more to say about these directions in the months ahead. We look forward to collaborating on these advances with you as our trusted customers and partners.

Sources

National Center for Education Statistics (2011). Common core of data public elementary/secondary school universe survey: School year 2009-10, version 1a [data file and codebook]. Retrieved from <http://nces.ed.gov/ccd/pubschuniv.asp>

Recommended citation:

University of Oregon, Center on Teaching and Learning (2012). *2012-2013 DIBELS Data System Update Part I: DIBELS Next Composite Score* (Technical Brief No. 1202). Eugene, OR: University of Oregon.