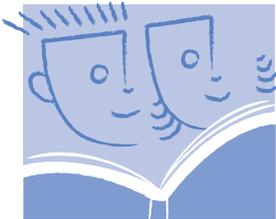# Position Paper on Use of DIBELS® for Student-Level Accountability Decisions

Ruth A. Kaminski and Roland H. Good III

Doris Baker, Kelli Cummings, Chantal Dufour-Martel, Kathleen Petersen, Stephanie Stollar, Kelly Powell-Smith, Joshua Wallin

Dynamic Measurement Group

DIBELS were developed and validated for the purposes of providing effective support to individual students by way of formative review of programs and student response to those programs. Because DIBELS were designed to provide information to help *support* students, it is not appropriate to use DIBELS scores for labeling, tracking, or grading students. Likewise, it is not appropriate to use DIBELS for retention and promotion decisions. Recent reports of the use of DIBELS for retention are especially alarming given the overwhelming information that has been published about the poor outcomes associated with retention. The NASP position statement on Student Grade Retention and Social Promotion is available at http://www.nasponline.org/about_nasp/pospaper_graderetent.aspx. Another excellent resource for scientific research on the effects of retention is http://www.education.ucsb.edu/jimerson/retention/.

The reasons DIBELS should not be used for high-stakes decisions for individual children are three-fold:

1. First, DIBELS were not validated for such uses. It is important to remember that although DIBELS have demonstrated technical adequacy for the purposes of screening and progress monitoring, they are one-minute measures administered at a single point in time. Many factors can impact a child's score and reliability of the measures is increased by repeated assessment over time (i.e., validating student need for support, progress monitoring within an Outcomes Driven Model of educational decision making. See the best practices chapter referenced at the end of this position paper for more information).

2. Second, DIBELS are deliberately intended *not* to assess a wide range of individual skills related to a domain. Instead, DIBELS were designed to be *indicators* of five key early literacy skills that are predictive of later reading achievement. When aggregated for an individual across time, the pattern of performance provides an important evaluation of the student's response to instruction for adjusting that instruction and providing additional support if necessary.

3. Finally, the use of *any* single indicator of competence to make important decisions, such as child retention, grading, or labeling, violates professional standards for educational measurement (AERA, 1999; APA 1999). The importance of using other relevant information, including multiple forms of assessment, and viewing assessment results within the context of the school cannot be overstated.

To use DIBELS data appropriately and effectively to support individual students, it is important that educators have a clear understanding of the conceptual and empirical foundations of the decision-making utility of DIBELS benchmark goals. The DIBELS goals and cut scores used for decision making are empirically-derived, criterion-referenced scores. The benchmark goals indicate the <u>probability</u> of achieving the next benchmark goal or the <u>probability</u> of the need for additional instructional support for the student to achieve the next goal. Thus, basing a student's grade or a high-stakes decision on a student's DIBELS score (i.e., his/her probability of achieving the next benchmark goal) would be inappropriate. Because the goals and cut scores are based on longitudinal predictive probabilities, they are not set in stone. A score at or above the benchmark, for example, merely indicates an

80% probability of achieving the next goal; but it is not a guarantee. Most students who meet a benchmark goal will need continued, high-quality instruction to hit the next target. However, the odds are that approximately 20% of students who achieve scores at or above the benchmark goal may still need supplemental support to achieve the next goal.

Scores that fall between the benchmark goal and the cutoff score for high risk represent patterns of performance where approximately 50% of students achieved subsequent literacy goals. Students with scores in this category require *strategic* planning on the part of educators to determine appropriate strategies to support the students to meet subsequent early literacy goals.

For students whose DIBELS scores fall below the cut point for high risk, the probability is low (i.e., < .20) of achieving subsequent goals without additional instructional support; yet these odds mean that a small percentage of students who score in this range may achieve subsequent benchmark goals anyway. The important use of this risk information is to provide a targeted, intensive intervention with frequent progress monitoring and formative modification of instruction to ruin the prediction of risk and ensure the student achieves subsequent benchmark goals.

We recommend that educators carefully consider the progress of students on all measures administered as well as all other information they have about the student as they evaluate their instruction and make educational decisions to support students.

DIBELS value is as a formative assessment tool to help teachers and administrators to identify needs for support. We recommend the following practices related to the use of DIBELS data as one piece of data to be used in making the following educational decisions:

a. Accurately identify need for support *early*,

b. Provide meaningful and important instructional goals,

c. Evaluate progress toward those goals,

d. Modify instruction as needed for students to make adequate progress,

e. Review outcomes quarterly.

For more information regarding the use of DIBELS for individual educational decisions, see "Best Practices in Using DIBELS in an Outcomes-Driven Model", available at http://www.dibels.org/pubs.html.