Understanding DIBELS® 8th Edition Composite and Measure Scores

The new composite score used in DIBELS 8th Edition is the most robust predictor of risk that DIBELS has ever offered. It is superior to any one subtest in its reliability and in the accuracy of its predictors. As a result, we advise that schools use the composite score in determining overall risk in reading.

The subtest risk statuses can be helpful in determining instructional groupings or understanding the underlying reasons for aggregate risk. Nonetheless, they do not overrule the composite status.

The composite score used in DIBELS 8th Edition was computed using data from thousands of real children. Using a confirmatory factor analysis (CFA), DIBELS researchers determined the weight to be given to each subtest. CFA takes into account each student's scores on all the relevant subtests and yields weights that can be used to calculate an overall score. The advantage of this approach is that it takes into account the different scales of scores for subtests, as well as the measures where students vary more in their scores. Importantly, these weights are based on real data and actual scores, *not on the risk status*, for each subtest.

Thus, the new composite scores give an empirically derived risk status that communicates each student's overall risk, taking into consideration performance on each subtest. As a result, it is possible for a student's risk status on one or more subtests to be higher or lower than the composite score risk status. More specifically, students on the cusp of risk or benchmark status (give or take a few points) across multiple subtests can end up with a composite that disagrees with the subtests. That difference is to be expected because students who are near a cut are, in essence, on one side or the other of a tipping point. These situations are generally infrequent but may be more frequent in some contexts depending on student performance.

When determining a student's overall risk, the composite score is the best predictor of end-of-year performance. Nonetheless, risk on each specific subtest remains important for determining where to focus individualized instruction and for assigning students to small groups. We provide below a few real-life illustrations of situations where composite status "disagrees" with subtest statuses to aid schools in interpreting these situations.

In most cases, a student scoring in a single risk category on all the individual subtests will score in the same category for the composite score; however, specific cases exist where the measure and composite colors may not align.

1. Green subtests with blue composite: Some students who score near the top of the green (at benchmark) range on all subtests may end up with a composite score in the blue (above benchmark) range. These students are doing very well and are at very low risk for future reading difficulties.

2. Yellow subtests with green composite: Similarly, students scoring near the top of the yellow (below benchmark) range on all individual measures may end up with a composite score in the green (at benchmark) range. Overall these students are performing well enough to be considered at low risk on the composite; however, we would recommend continuing to attend to these students' individual skills to ensure they receive adequate instruction to meet their needs.

3. Green subtests with yellow composite: Students scoring near the lower end of the green (at benchmark) range on all measures may end up with a composite score in the yellow (below benchmark) range. Although these students have reached the benchmark cut point on each measure, their lower-end scores across multiple skills may result in a composite score that shows they could be at some risk for future reading difficulties. We would recommend continuing to provide core instruction but keeping a close eye on individual skills and providing the necessary reinforcement to ensure these students remain in the green (at benchmark) range throughout the year.

4. Yellow subtests with red composite: Similarly, students who score at the lower end of the yellow (below benchmark) range on subtests may end up with a composite score in the red (well below benchmark) range.

5. Varying subtest scores: In cases where students score at differing levels on subtests, the composite will not necessarily reflect an average of the subtest levels because the composite is not based on risk status. Instead, the more informative subtests will be given more weight. For example, if a student scores green on three subtests, yellow on one, and red on another, it matters which subtests are at which levels. If the student is in second or third grade and received the red status on ORF-WRC, it is much more likely that student will obtain a yellow or even red composite score. In contrast, if that same student received the green status on NWF-CLS, WRF, and ORF-WRC, it is much more likely the student will obtain a green composite score, Importantly, even though the latter instance indicates a student is at benchmark, a red status on either NWF-WRC or ORF-ACC may be cause for concern and monitoring depending on how deep in the red the student's score is.

In the coming months, we will provide more concrete examples in multiple grades with implications for instruction to help guide schools and teachers further in their decision-making.

