



A Community of Learners

Informational Memo - Computer Adaptive Assessment Selection

TO: School Board
Superintendent Kocanda

FROM: The Computer Adaptive Assessment Committee
Phil Earvolino, District Data Consultant

DATE: March 15, 2016

Overview & Background

The Computer Adaptive Assessment Selection Committee is continuing its work to determine which of the adaptive assessments will be the right fit for Winnetka District 36. Using the following Essential Questions to guide its work, the committee discussed the purpose of standardized assessments and the needs of the district.

- Which computer adaptive assessment is the best fit for D36?
- How can computer adaptive, nationally-normed tests provide D36 with feedback on student progress?
- How do we use the data we collect from computer adaptive tests?

We reviewed the purposes of standardized assessments:

- Provide data about individual student needs, progress and achievement.
- Serve as a universal screener to determine if further diagnostics are warranted in an effort to provide targeted interventions.
- Collect a record of longitudinal growth in a time efficient manner.
- Compare local student growth and achievement to national student growth and achievement.
- Determine the strengths and areas in need of improvement within the local curriculum.

As we gathered information about the NWEA MAP and Scantron assessments, we polled staff members to ask them to share their experience and thoughts, we contacted other

districts that use these measures, met with the District's data analysis consultants to discuss assessment options, and scheduled a webinar with representatives from each assessment team.

March 2016 Update

We found that there are many similarities between NWEA MAP and Scantron Performance Series. Both assessments are widely used nationally and have the following capabilities:

- can monitor student, school and district performance over time;
- can help evaluate curriculum;
- can be used as universal screeners for Response to Intervention requirements;
- offer “add-on” programs to monitor the progress of students who receive support;
- offer parent resources;
- offer professional learning opportunities.

Staff members who have had experience with the assessments have shared that each assessment provides reports that teachers and administrators can use immediately to learn about their students' progress to adjust instruction as necessary. Both offer information for instructional groupings and differentiation that can assist teachers when planning lessons.

The differences between the two assessments are:

- NWEA parent resources are available online and can be made available for parents electronically while Scantron offers printed parent materials.
- Scantron provides growth data based on national data while NWEA can provide [Virtual Comparison Groups \(VCGs\) reports](#) which show how students are growing compared to similar students educated in similar schools across the country.
- Data models used by the companies differ.
- [This chart](#) goes into further detail about the differences between the two assessments.

We followed up with NWEA MAP to learn more about their research and to dig more deeply into the quality of their data. The District's Data Analyst, Philip Earvolino, participated in this discussion. Mr. Earvolino has experience with the data collection methods used for Scantron Performance Series and can provide a knowledgeable comparison of the methods used. The results of Phil's research on how each company gathers and uses data follows:

Data Analysis: Phil Earvolino's Perspective

Scantron Performance Series and MAP assessments are both computerized adaptive tests, as was STAR. For all of these assessments I see no significant problems with 'ceiling' effects, i.e. a student reaching a perfect score. While it does happen, particularly for seventh and eighth graders, it is quite rare, even for advanced students.

Scantron Performance Series

For each grade and content area, projected spring scores are specified as follows: based on Scantron's national norm group, fall scores are broken into deciles. For each decile, a spring target score along with an associated standard deviation are provided. Thus, there is no growth model, per se, but a "map" between, effectively, ten fall scores and their matched spring scores. The problem with this approach, particularly as implemented in a high achievement school district, is that it is too generalized, especially at the tails. Hence, a student scoring in the 90th percentile nationally in the fall is projected to have the same nominal growth (i.e., the spring score minus the fall score) as a student in the 99th percentile. However, based on my own research, the actual observed growth scores for these two students may be considerably different. The same observation is relevant to students on the low side of the achievement spectrum. Furthermore, the Performance Series assessment only projects fall-to-spring scores (their fall-to-winter scores are merely the fall-to-spring scores divided approximately in half).

NWEA MAP

MAP employs a fully-developed model incorporating three years of test data from over ten million students per cohort. The model is "mixed level", which means it examines student variation, as well as school district variation (essentially treating each school district as a single test-taker). The model also incorporates instructional time (specifically, the actual number of school days between consecutive tests, such as those offered in fall and spring, or even fall sessions in consecutive school years). In this regard it is unique among the major vendors we have examined. For a given student, a student's growth (whether from fall-to-spring, fall-to-fall, or other pre-test/post-test intervals) is expressed as a "conditional growth" percentile or index. The former is similar to the SGP value generated by Renaissance for use in STAR testing: it ranks the student's actual growth on a scale of 1 to 99 as compared to that of other like students in the nationally normed sample. The conditional growth index expresses this growth in standard deviations above or below the norm group mean. With the MAP assessment the projected spring score for each student is based on several of his or her previous scores, not just from the fall window, but from sessions in a previous school year. By

incorporating more than one previous score, the model is much more stable and robust than those which look only at a single previous score.

MAP also provides what it calls Virtual Comparison Groups, by which it claims to match the scores of any given student with those of up to fifty students in similar school districts in NWEA's database. However, it determines similarity on the basis of, apparently, only two criteria: percentage of students eligible for free or reduced lunch and school district community setting (i.e, rural, urban, or suburban). This will not help us get what we are really looking for, which is a comparison to students in other affluent high-achievement districts. From what I've seen, though, no one can provide that because it would require integration with, e.g. Census data.

Next Steps

Based on the differences in the research and data models, the evidence is most aligning with NWEA MAP to be used as an assessment for District 36. Next Steps include scheduling time for NWEA representatives to meet with other key stakeholders, including representatives from classroom, special education, technology, and teachers on the PERA committee to determine if they see any additional benefits or limitations of this assessment.

The committee plans to bring its final recommendation to the Board at the April meeting, potentially using Work Session time have deeper conversations about the decision.