REL Mid-Atlantic Webinar
Teacher Evaluation: Alternate Measures of Student Growth
Q&A with Brian Gill
October 29, 2013

This webinar described the findings of our literature review on alternative measures of student growth that are used in teacher evaluation. The review focused on two types of alternative growth measures: statistical growth/value-added models and teacher-developed student learning objectives. This Q&A addressed the questions participants had for Brian Gill following the webinar. The webinar recording and PowerPoint presentation are also available.

General questions

1. **What assessments does the research suggest are most effective in measuring student growth?**

   When examining assessments for measuring student growth, we considered the validity and reliability of the growth measures. The literature suggests that value-added measures based on low- stakes commercially available assessments, such as the Stanford Achievement Test, correlate moderately with other performance measures. Studies have also found that the reliability of value- added measures based on these types of assessments is comparable to that obtained using state assessments. Less is known about the validity and reliability of course-specific assessments, including locally developed tests, but the evidence indicates that these types of assessments have validity and reliability comparable to other performance measures. Validity and reliability of growth measures in teacher evaluation will depend in part on the consistency of the assessment across teachers, on the alignment of the assessment with the curriculum, and on the sophistication of the statistical method used to attribute growth to teachers.

2. **The real value in using value added models (VAMs) or SLOs [student learning objectives] is what it tells the teacher in terms of how to improve instruction. What research has been done in this area?**

   This is an area where research is just getting started. Even in the absence of a lot of rigorous research, school districts and schools can look themselves at the SLO results in the context of information about professional practice. There's enough professional practice knowledge among educators that it ought to be possible to use data from SLOs to help in coaching teachers to improve instruction.

3. **If you are using an RTI [response to intervention] framework and doing universal screening several times per year, might that be a good way to measure student growth schoolwide? For individual teachers?**

   We don't know enough about response to intervention to know if this would be a good way to measure schoolwide student growth. It is possible that the RTI framework can be used to measure student growth. Universal screening is certainly something that would promote reliability and validity, and there might be ways to apply value-added models to the screening results.

4. **How do you work with teacher associations on new evaluation methods, and how successful have you been working with them?**

   Some districts have successfully engaged teacher associations in the design and implementation of new evaluation methods, including new measures of professional

practice as well as student growth measures. Teacher associations may be more interested in participating actively in the development of new evaluations if the evaluations are embedded in a system of supports designed to help teachers improve their practice.

5. Were the studies reviewed primarily for elementary grades?

No. Commercially available assessments are often used in elementary and middle school grades, but course-specific assessments are typically used in secondary grades.

Questions about value-added models

1. How do you recommend communicating the uncertainty in these value-added estimates?

A key point with regard to uncertainty is to recognize that all measures have some degree of uncertainty. Principal observations of teachers also have uncertainty, but it is not always recognized because we don't quantify it in the same way as we do for value-added measures.

The aim is to try to reduce uncertainty as much as possible. With value-added models, this can be done by averaging across multiple years of data and by applying statistical shrinkage adjustments. There is evidence suggesting that these approaches reduce the level of uncertainty in value-added measures to a level comparable to other measures of job performance that are used outside of teaching.

It will be important to explain to stakeholders that there is uncertainty in all measures of teacher performance and that the district is taking specific steps to reduce this uncertainty as much as possible.

2. What, if any, states are considering using attendance in VAMs?

We don't know of any states that are considering using attendance for teacher evaluation purposes. Pittsburgh is interested in using attendance and other types of nontest outcomes in value-added models to evaluate school performance, but not teacher performance.

3. Are value-added models being used to evaluate special education teachers?

To our knowledge, applying value-added models to special education teachers is relatively rare. One concern is that assessments may not be able to measure achievement and changes in achievement for special education students in the same way that they can for other students. Another potential issue is that there may not be a large enough number of special education teachers to estimate value-added models. This is an area that requires more research.

Questions about student learning objectives

1. Are there common potential problems which should be anticipated—and addressed—prior to beginning the development of Student Learning Outcomes?

Districts should be aware that teachers and evaluators (typically principals) may have to go through substantial training on how to develop and use SLOs. Training is important so that teachers select or develop appropriate assessments, create challenging and meaningful growth targets, and understand how to use the SLO data to improve teaching practice. Evaluator training is also important to ensure that SLOs are consistently applied across teachers and schools. Districts may want to consider implementing a centralized system to audit and approve SLOs to improve consistency, but they should be aware that this creates

additional burden at the central office. Whether done by evaluators or district staff, ensuring consistency is a key issue because it will help address concerns about fairness that have arisen in some districts, particularly if SLOs are being used for evaluations or high-stakes decisions.

Another implementation issue to be aware of is that both teachers and evaluators may have to spend substantial time developing, approving, and assessing SLOs.

Lastly, districts should be aware that SLOs can be useful tools for planning instruction, understanding students' needs and abilities, and improving teaching practice over time. Using SLOs for evaluating teachers may undermine their value for the purposes of instructional planning and improvement because there may be a conflict of interest between setting achievable goals for evaluation purposes and setting challenging goals for instructional purposes.

2. **Are you seeing more SLO tasks that are performance-based or pencil and paper tests?**

Both types of assessments are used, and use varies to some degree by grade and subject. We are not aware of a systematic catalogue of the types of assessments being used in SLOs.

3. **How should performance measure data be collected from teachers?**

The way in which data are collected varies depending on the type of performance measure.

4. **How can a teacher ensure that the student population is correct for SLO/SGO [student growth objective] purposes? With kids coming in and out of the classroom, how long does a student need to be in the class for them to count towards a teacher's SGO rating?**

In many cases, teachers will be conducting the assessments used for SLOs/SGOs, so they will have the ability to ensure that students are included only if they have spent substantial time in the class. Districts and states should set attendance thresholds for SLO inclusion in order to ensure consistency across teachers and schools. Many districts and states already have such rules for calculating value-added, and they could use the same thresholds for SLOs/SGOs.

5. **How should student ability (special education vs. gifted) be factored into student achievement data for teacher evaluation?**

Value-added models consider students' prior achievement, and sometimes also their special education or gifted status, when predicting student achievement growth over the course of a school year. Similarly, SLOs are also determined based on students' achievement on a pretest. It would be appropriate for districts to allow teachers to adjust SLOs for some students based on student characteristics such as special education status.

6. **I have some validity concerns, especially with growth measures in the arts.**

Concern about the validity of growth measures used in arts education is not unreasonable, given the lack of evidence about such measures. Unfortunately, we have not yet seen any research that examines the validity of SLOs/SGOs in the arts.

7. **Where can we find resources for creating SLOs (best practices)?**

Here are three good places to start:
- Community Training and Assistance Center: http://ctacusa.com/studentlearningobjectives.html. This group has done a lot of work on implementation of SLOs and professional development and also some research on SLOs.

- Center on Great Teachers and Leaders: http://www.gtlcenter.org/learning-hub/student- learning-objectives.

- Pennsylvania recently posted an online SLO training website: http://pdesas.org. Click on the "Homeroom" button. Follow this link and register on the Research in Action website—this is an open resource.

8. What weight do SLOs typically have in evaluations?

The weight of SLOs in evaluations is dependent on state and district policy. In some states, districts have a lot of local control in determining these weights. When SLOs are used along with value-added measures to extend coverage beyond tested grades and subjects, a fairly typical practice is to assign SLOs the same weight that value-added is assigned for teachers in tested grades and subjects.

Alternative measures used by webinar attendees' districts/schools

- NYS [New York State] is allowing some interesting space for districts to experiment with tasks!

- You've already mentioned it, but in Pittsburgh we use CBAs [computer-based assessments]— tests developed by the district. I think that there is a push to see how SLOs can be incorporated (I could be wrong).

- District developed assessments, nationally recognized assessments, industry certified exams, student projects, and student portfolios.

- Teachers selecting CCSS [Common Core State Standards] that have been identified as essential to success in the course in HS math. Using a pre and post assessment that addresses this CCSS.

- NWEA [Northwest Evaluation Association] using SGPs [student growth percentiles] is being tested in an elementary school district in Arizona.

- Instructionally sensitive summative learning targets (district assessment); DBQ (document based questions) with department created rubric

- Also using NWEA MAP [Measures of Academic Progress]

Responses to the question, "As a result of today's webinar, what action steps do you plan to take?"

- Hoping that PARCC [Partnership for Assessment of Readiness for College and Careers] and data collected by the NJ SMART [New Jersey Standards Measurement and Resource for Teaching] will provide more support to the process. Looking for support in determining a more reliable process. Glad to see we are all struggling together.

- Explore the resources that were provided as well as discuss with our SLO team the next steps for training our districts.

- I truly appreciate the resources provided.

- Look to review the SGPs created and develop a set of lessons learned and steps to follow for next year's SGO development.

- Training and piloting the SLO development process and providing resources and

- training around assessment literacy.
- Appreciate the resources from colleagues.