REL MID-ATLANTIC Regional Educational Laboratory at ICF International
Bridge Events and Webinars
Supporting Education through Research

REL Mid-Atlantic Teacher Effectiveness Webinar Series
Using Student Surveys to Monitor Teacher Effectiveness
Q&A with Dr. Ronald Ferguson
June 27, 2013

In this webinar, Dr. Ronald Ferguson, creator of the Tripod Project and Senior Lecturer at Harvard University Graduate School of Education, discussed the use of student surveys as an approach to measuring teacher effectiveness. This Q&A addressed the questions participants had for Dr. Ferguson following the webinar. The webinar recording and PowerPoint presentation.

Survey Development and Administration

1. **What are the different types of surveys that can be used to monitor teachers' effectiveness?**

Most of the surveys administered in schools are whole-school climate surveys, not classroom-level surveys. In order to reach judgments about individual teachers and individual classrooms, we need to ask questions about individual teachers and classrooms not whole schools. Also, it can take a long time to refine a survey. We have been refining the Tripod surveys for more than a decade. When the Bill and Melinda Gates Foundation Measures of Effective Teaching (MET) project came along, the Tripod survey was the only classroom-level survey found that had been widely used and refined sufficiently for use in that project. A few other classroom-level surveys are now under development or in early phases of use.

2. **Do you view student surveys as more of a professional development tool or a diagnostic tool?**

Both. I see diagnostic and professional development uses as linked—part of the same process. Student perceptions of instructional practices can help identify strengths and weaknesses in teaching. They provide information on the basis of which teachers and their supervisors can diagnose problems and set specific priorities for differentiated professional development and coaching support. In addition, surveys can be used above the classroom level to identify patterns across clusters of classrooms, schools, or even districts. Beyond instructional practices, per se, student surveys can provide evidence about student engagement. An interesting and useful professional development activity is to consider how students' perceptions of teaching (for example, clarity of explanations) and engagement (for example, help-seeking behaviors) are related.

3. **What is the relevance/usefulness of student surveys at early grades?**

Even young children have points of view about how things are going in their classrooms, and teachers find those useful. For example, teachers can review their personal classroom reports on student responses and consider whether their students' perceptions match their own and, if not, what the implications might be. I am reminded of a kindergarten teacher who considered it very useful recently to know that only half of her students agreed that the

class stays busy and does not waste time. Districts that plan to use student surveys as part of their teacher evaluation system usually include elementary-level surveys in the evaluation. It is understandable that people are surprised young children can respond to surveys with good reliability for distinguishing between classrooms. Nonetheless, we find very informative patterns in student responses at the elementary level, even for the early grades.

4. **How can I get teacher buy-in?**

   A few suggestions for ways to increase teacher buy-in include (1) explaining how the implementation methods will protect their confidentiality; (2) describing some of the ways that survey results can be used to help teachers improve; (3) committing to providing differentiated professional development that responds to survey results; and (4) explaining that the survey is a worthwhile use of classroom time because it can improve the quality of class time by illuminating student thinking in ways that improve teaching and learning.

5. **How do you increase responsiveness to student surveys?**

   This does not tend to be a problem. Responsiveness is usually very high. Although there can be a slight fall off from the start to the finish of the survey, overall response rates are typically over 90percent because surveys are administered to students during class time. Students are given the option to skip any item they don't understand, but most don't skip very many items. The most common time there is an issue occurs when students in secondary schools are asked to complete the survey for several classes within a day or two. If students are likely to see the survey in multiple classes, it is best to stretch the survey period out over a couple of weeks.

6. **How do you separate nonsense student surveys ("Bad teacher" = too much homework) from more sensible surveys?**

   The items on the student surveys do not ask students what constitutes an effective teacher or whether they like their teacher. Instead, the items ask for levels of agreement with descriptive indicators of what is happening in the classroom. We have found that students actually enjoy this activity because they are not usually asked to provide this kind of input. There is seldom evidence of inauthentic responses. Nevertheless, we recognize that this is a real concern. We have ways of checking patterns of classroom data to discern if students are taking the survey seriously (e.g., too much uniformity across a class's responses or if negatively worded answers are not systematically the opposite of positively worded answers, duplicates).

Validity and Reliability

7. **How reliable and valid are these student surveys? What methods were used to assess the effectiveness and validity of the surveys?**

   There are various statistical ways of rating a survey for reliability. We favor multilevel reliability measures that take into account levels of agreement among the items in multi-item indices for particular concepts and also among the students in classrooms. We know a survey is reliable when there is a high level of agreement between different survey items that are supposed to measure a particular concept and also among the students in the class. The

Tripod surveys have been refined over time to achieve high levels of reliability. Another check on reliability is to examine the correlation between survey measures collected from different classes taught by the same teacher. The MET project concluded that the between-classroom correlation between composite measures of teaching quality is higher for student surveys than for both value-added test score gains and adult observations of teaching practice.

The primary way to document validity is to show that student surveys are correlated with other measures of teaching quality taken from the same classrooms. Showing this was a major contribution of the MET project. Six districts were involved in the MET project. Teachers were chosen in groups of two or three in each school so they could be randomly assigned to classrooms. A random assignment evaluation of the relationships between observational rubrics, student survey responses, and value-added achievement gains was conducted. The fact that teachers were randomly assigned to classrooms helps to validate the conclusions reached concerning the validity of the different metrics as measures of teaching quality because each measure still tended to predict the others, even after randomization. The MET project reports can be found online at www.METProject.org.

8. **Can you comment on the issue of survey anonymity, and do you have any findings that suggest effects (or lack thereof) of anonymity on validity?**

The surveys are anonymous at the student level. Paper surveys have a peel-off label with the student's name and a barcode on the survey to match student information to other information. Online surveys are similarly anonymous. No information that is identifiable at the student level is returned to the school. In addition, at least 5 and sometimes 10 students are required before any classroom level data are reported to the teacher.

At the teacher level, results from the first administration of Tripod surveys are often anonymous in the first year. In these cases, only the teacher receives a report that can be linked to a specific teacher or classroom. However, when student survey responses become part of teacher evaluation formulas, someone in the system will of course need to have identifiable information. At that time, it is important to have processes in place to manage the information properly, sharing it only on a need-to-know basis. Whether there is any effect on validity is something we will need to track over time.

9. **Are there ways that results of a student survey can be interpreted as a measure of student achievement?**

Not exactly. The engagement measures and Tripod's 7Cs framework of effective teaching that we measure using student surveys help us to understand what happens in classrooms. The 7Cs framework captures much of what people talk about when they describe effective teaching (Caring, Captivating, Conferring, Clarifying, Consolidating, Challenging, and Controlling).The combination of what teachers do and what students do produces student achievement outcomes. So, data for these indicators (e.g., the 7Cs framework and student engagement targets) can provide insight into what produces student achievement outcomes, but they are not themselves measures of achievement. If measures of achievement are not available, they might be considered reasonable proxies, but only proxies.

Educator Effectiveness

10. What specific data do the survey results provide, and how does it align with educator effectiveness requirements?

The survey results provide data about student engagement metrics and the learning environment from the Tripod's 7Cs indices (Caring, Captivating, Conferring, Clarifying, Consolidating, Challenging, and Controlling). We can rank order classrooms on how students respond on the 7Cs measures. Specifically, we can take all the responses for any item and put together an average for the entire class. If we put together all the 7Cs indices and all the responses, we can come up with a single summary number on how the students rate that classroom on the 7Cs framework. Using the summary numbers, we can rank order all the classrooms in any particular set of data. The objective is to use this information in order to focus and differentiate professional learning opportunities for teachers. Most places are using value-added and observational measures like Charlotte Danielson's Framework for Teaching to make judgments about educator effectiveness. The MET project established that Tripod surveys correlate with both value-added and observational measures and tend to be more stable than either as a measure of classroom-level effectiveness for the grades and subjects studied by MET.

11. How do you see student surveys being part of the new teacher evaluation systems based on the Danielson framework?

There is a good degree of alignment, and there are useful ways of using the frameworks together. The Tripod student surveys are a bit more reliable, but each instrument does have informational value the other does not, so bringing them together is beneficial. I am working on a paper about the alignment now, which will be part of an edited volume of work by contributors to the MET project.

12. How much weight if any will be applied to student surveys in teacher evaluations?

There is no right answer; it depends on the criteria one is using to make the decision. One criterion is reliability; on this dimension, Tripod surveys rate strongly. Another is validity, and here again the Tripod surveys are strong. On the criterion of receptivity, teachers might be more receptive to observational frameworks where the rating came from an adult. Also, it will matter how useful the ratings are for use in coaching discussions. Thus, the metrics, the way the metrics complement each other, and the social processes of adult interaction when using the metrics can all influence decisions. As an example, Memphis' teacher evaluation measure includes student value-added, scores from supervisors or others observing classrooms, and Tripod student survey ratings; the student surveys are 5 percent of the evaluation. Another district has discussed using 15 percent but is still deliberating. I think 15 percent should probably be the upper limit for now.

Creation of Student Surveys

13. What are the guidelines to create valid surveys for the purpose of getting student information?

Most experts recommend that you use a survey others have created and validated rather than try to create a survey yourself, especially if you are going to use it for any official purpose. The Tripod surveys were created over 13 years and continue to be refined each

year. If you care about having a high-quality survey, you will want to find one that has been refined over a period of time. Nevertheless, the first general step is to be conceptually clear on what you want to measure. The second step is to develop or have multiple items to capture each concept. The third step is to do cognitive testing where you sit down with the types of people to whom you will be administering the survey and have them interpret what the survey items mean to them. If their interpretations for a particular item differ from what you intended, you need to change or replace the item. The fourth step is to administer the survey in a pilot and confirm that the results are interrelated as predicted (e.g., items that are supposed to measure the same concept turn out to be highly correlated in responses). Typically, you will find some items do not behave as expected, and you will need to develop new items, and even the new items may or may not work, so you will probably need to go through several rounds. In the meantime, you may have refined your theoretical conceptions, so you will need to modify your survey to stay updated. This process may continue for years until it converges.

14. **How would you create a document or a tracking system to evaluate student surveys?**

    I assume this question concerns documents for tracking and evaluating the survey implementation process. For example, making sure that identifiers are correct and that surveys are labeled, delivered, and collected from the right students, classrooms, and teachers. This is a specialized set of functions. I do not know the details. For Tripod surveys, these functions are handled by Cambridge Education, which operates Tripod and delivers services related to data collection and standard analysis and reporting. See TripodProject.org for more information.

15. **How can I design, collect, and analyze data efficiently? When and how are data collected (paper, electronic)?**

    Use multiple measures, deployed multiple times per year, over multiple years to accumulate information and reach more reliable judgments about what is happening in classrooms and identify trends. Paper is more common due to the often limited online capacity at schools.

16. **How do you write and vet effective questions?**

    Please see responses to related questions above, especially **"What are the guidelines to create valid surveys for the purpose of getting student information?"**

Application to Practice

17. **Would student surveys yield meaningful results in the area of special education (e.g., education specialists)?**

    The surveys are general, but the answer here depends on whether students are able to comprehend what they are being asked (i.e., the students' developmental levels). We have surveyed lots of special education students in the past, and the patterns in their responses make sense. Still, there may be students who are sufficiently delayed developmentally that the age-appropriate survey is not developmentally appropriate. Similarly, for English language learners capacity to understand the items is important. The Tripod surveys are available in Spanish and eight other languages, but other languages may require active translation. The survey can be read to a student in his/her native language, and the translator can make sure that students know how to check the appropriate responses.

18. **What setting is ideal for administering such surveys, with the teacher in question or**

with another teacher?

There are several considerations. Typically our initial impression, and this still has some weight to it, is that it should be someone other than the teacher so students feel certain that the teacher will not see their responses. However, a consideration that weights toward having the teacher in the classroom is that the teacher might be the best at keeping students on task, and we want students to consider that teacher as they complete the survey. In either case, for Tripod, the students are given peel-and-stick envelopes in which to insert their surveys before handing them to the teacher or the person monitoring the completion of the surveys. The win, win answer is to have a teacher and third party in the room to monitor. The teacher is there to keep students on task and to ensure students relate their responses to that teacher, but a third party is there to oversee the integrity of the process.

Action Steps

Participants responded to the question "As a result of today's webinar, what action steps do you plan to take?" Some of their responses are listed below.

- As a state administrator, I plan to press harder for the use of student surveys multiple times over multiple years.
- As a PLC team leader in my school, this has given me food for thought. This is something I want to bring to my team.
- We've surveyed students for many years, but the questions are school-generated and have not been proven reliable as indicators of student success. I'd like us to use more valid questioning in student surveys.
- I plan to better understand the 7Cs; share information with relevant stakeholders; get buy-in; and then use the 7Cs as a framework for transformation.
- We have had Best Practices sharing seminars, because our teachers have been expressing the desire to hear from live examples of success. They don't want an outsider to come in to tell them how it is done. So I think surveys in this instance will be very telling. We want the teachers with high student ratings to share their practices.
- We have been refining a survey we used for several years, and it's encouraging to hear that has been part of this more robust process. I'm going to continue to advocate for the use of student surveys as part of our feedback and continue to refine them to align with our own frameworks for teacher effectiveness (Danielson).
- I think the slide on differences in months of learning per school year between the 25th and 75th percentile classrooms is a simple yet powerful way to illustrate the importance of student surveys. It could be an important part of communication efforts for school planning to implement student response measures.

Additional Resources

- Article in *The Atlantic* about the Tripod Surveys: Why Kids Should Grade Teachers
- The Tripod Project
- Cambridge Education
- Measures of Effective Teaching Project