EDUCATION

TEACHER *and* LEADER
EFFECTIVENESS (TLE)

<u>**Value-Added Measures: Frequently Asked Questions**</u>

Value-added overview

Q: What is value added?

A: "Value added" is a measure of a teacher's contributions to students' academic growth. It is one of multiple measures that will be included in Oklahoma's Teacher and Leader Effectiveness (TLE) system.

To estimate a teacher's value-added result, a value-added model compares two sets of test scores: (1) the average actual scores that the teacher's students earned and (2) the average scores achieved by the students' "typical peers" throughout the state. The difference between these two sets of scores is the teacher's value added.

In Oklahoma's value-added model, "typical-peer scores" are estimated by looking at the achievement of students' most similar "peers" in the state. These peers are similar in terms of scores earned on multiple assessments and other background characteristics.

It is important to note that by looking specifically at "typical-peer scores," the value-added model is designed to isolate a teacher's value added from other factors that might affect a student's scores but that are outside the teacher's control. These factors include students' status as English-language learners, use of individual education plans, and attendance at school.

Q: What are the "pre-test" and "post-test" scores used in the value-added model?

A: Educators may hear the terms "pre-test" and "post-test" used to describe how value-added results are calculated. In this context, the post-test is the state assessment taken by students after they have been taught for the majority of the year by the teacher being evaluated. Pre-tests are the tests taken by these same students near the end of the prior school year, usually before they were taught by the teacher being evaluated.

Although the pre- and post-test scores are important, the value-added model does not directly compare these scores to determine a teacher's value added. Instead, the model compares a student's actual post-test score and the estimated typical-peer score on the same test. The typical-peer score is estimated based on pre-tests in multiple subjects and other student characteristics.

Q: *How is a student's typical-peer score calculated? Is there a formula?*

A: A student's typical-peer score reflects the average post-test scores of the student's most similar "peers" in the state. The typical-peer score is estimated using a statistical method that relates students' post-test scores to their pre-test scores on several assessments as well as additional student background characteristics. Because the statistical model is estimated after students take the post-test, there is no predetermined formula used to calculate typical-peer scores. Oklahoma is partnering with Urban Policy Development and Mathematica Policy Research to estimate typical-peer scores and value-added results.

Q: Is it possible for a teacher to receive a negative value-added result? If so, how does this impact their overall evaluation?

A: Yes. Since a teacher's overall value-added result represents how their students performed on average compared to similar peers, a negative value-added result simply means that their students performed below the average achievement level of their typical peers. If a teacher's students score at or above the average achievement levels of their typical peers, their value-added result will be at or above zero.

Although the value-added result itself can be a positive or negative number, this is not true of the value-added TLE component score that is included in the overall effectiveness rating. All value-added results are converted to a TLE component score of 1-5. Therefore, even if their value-added result is negative, it is not possible for a teacher to receive a value-added TLE component score of less than 1.

Q: What other states and districts has Mathematica worked with to develop value-added models?

A: Mathematica has worked with a number of states and districts to develop and customize value-added models, including Pittsburgh Public Schools; the Pennsylvania State Department of Education; Memphis Public Schools; District of Columbia Public Schools; and the Office of the State Superintendent in Washington, DC.

Q: How can we compare students' performance on two different tests and get accurate value-added results for teachers?

A: The Oklahoma state assessments are not designed to allow direct comparisons of students' numerical scores from grade to grade. A teacher's value added therefore cannot be calculated by simply subtracting the post-test score from the pre-test score. Instead, we estimate value added by comparing students' scores on a post-test and estimated typical-peer scores on the same test.

For example, if a teacher's students scored an average of 800 on a pre-test and 810 on a post-test, the teacher's value-added result would not be 10. Rather, we would compare the average post-test score of 810 with the average typical-peer scores for those students. In a very simple value-added model that estimated typical-peer scores based on only a single pre-test, the model would examine the performance of the students' peers—other students in the state who scored 800 on the pre-test. If the peers' post-test scores averaged, say, 790, the teacher's value-added result would be 20 (810 minus 790).

For additional accuracy, the typical-peer scores are based on multiple prior assessments and additional student characteristics. Incorporating this additional information can lead to more accurate typical-peer scores, based on peers who are very similar to a given teacher's students.

Q: How will Oklahoma's value-added model address factors that might affect student scores but are outside a teacher's control? Will these factors change over time?

A: The value-added model will adjust for a range of student characteristics that are outside a teacher's control. For the two-year pilot, the model will estimate typical-peer scores based on the following student characteristics: prior achievement on multiple assessments, status as an English-language learner, use of individual education plans, mobility across schools during the school year, attendance during the prior year, race/ethnicity, gender, and eligibility for free or reduced-price lunch. As these characteristics may influence student test scores, the model will separate the effects of these factors from the teacher's contribution. The characteristics used in the model will be re-evaluated after the pilot and may change as more data are gathered.

Q: Is it possible to include other factors, such as level of parent involvement or other environmental conditions that might impact a students' academic performance?

A: The value-added model can account for student characteristics that are measured accurately in the state's data system. Some factors that the state's data system does not directly measure may be related to student characteristics that are included in the value-added model. If so, then the typical scores can reflect these factors even though they are not directly included in the model. For example, parental involvement is likely to be captured in part by including students' prior achievement in the value-added model.

Q: Previously, teachers in the state have simply compared their students' pre-test and post-test scores. Isn't this an easier way of assessing student academic growth?

A: The Oklahoma state assessments are not designed to allow direct comparisons of students' numerical scores from grade to grade. But even if they were, the value-added model provides a more accurate reflection of a teacher's performance than a simple comparison of pre- and post-test scores. This is because the value-added model separates each teacher's contribution to student test scores from other factors that may affect the scores.

To isolate teachers' contributions, the value-added model accounts for the background characteristics of each student, including his or her prior scores on multiple assessments. Numerous studies have shown that such characteristics are linked to students' academic progress over the year. A simple comparison of student test scores from year to year would not adjust for these characteristics and would therefore reflect factors beyond teachers' control, such as students' status as English-language learners, use of individual education plans, and attendance at school. The Oklahoma value-added model adjusts for these factors by comparing students' actual scores to typical-peer scores.

Q: Does the value-added model account for learning loss over the summer?

A: Yes, the value-added model automatically adjusts for summer learning loss to the extent that the loss experienced by a teacher's students is similar to the loss experienced by the students used to estimate typical-peer scores. Because the model uses a large number of student characteristics to estimate the typical-peer scores, it allows for a great deal of variation in summer learning loss from one group of students to another.

Q: Can the value-added model show growth for students who start with very high scores?

A: Yes, effective teachers can achieve high value-added results regardless of their students' pre-test scores. This might seem surprising given that students who scored very well on pre-tests have little room to improve on their post-tests. However, value-added models adjust for this issue. Students who score well on one test tend to also do well on the next. However, on average, the students with the highest pre-test scores score slightly lower on the post-test. This means that a student scoring at the top of the pre-test scale will generally have a typical-peer score below the top of the post-test scale, leaving room for improvement. Thus, an effective teacher with high-scoring students on the pre-test can still have high value-added result if his or her students outperform their typical peers.

Q: What research has been done to demonstrate the validity of value added measures?

A: The Measures of Effective Teaching (MET) Project conducted one of the most comprehensive research studies on teacher effectiveness from 2009 to 2012. Their research highlights some key findings about value-added measures,

including measures of the relationship between value-added results and other teacher effectiveness measures, such as observations. Links to several key research papers the MET project produced are provided below.

http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf(full version)
http://www.metproject.org/downloads/MET_Gathering_Feedback_Practioner_Brief.pdf(brief version- summary of key findings)
http://www.metproject.org/downloads/MET_Validating_Using_Random_Assignment_Research_Paper.pdf (this paper describes a study to test the validity of value added using random assignment)

Another significant research study is the 2011 National Bureau of Economic Research working paper by researchers at Harvard and Columbia, _Long Term Impact of Teachers: Teacher Value-Added and Student Outcomes in Adulthood,_ which compared value-added estimates to student outcomes in adulthood and found strong correlation between them. Below is a link to a _New York Times_ article summarizing the results of the study and highlighting key findings. Please note that while the findings are relevant to our work in determining the best value-added approach for Oklahoma, the specific recommendations of the authors about how to use value-added results may not reflect the way they will be incorporated into evaluations in Oklahoma. Value-added results in Oklahoma's TLE system are one of multiple measures of effectiveness that will be used to support continuous improvement.
http://www.nytimes.com/2012/01/06/education/big-study-links-good-teachers-to-lasting-gain.html

Q: What states are using value-added measures for teacher evaluation?

A: Some of the states and large school districts using value-added measures currently include: Charlotte-Mecklenburg, the District of Columbia Public Schools, Florida, Los Angeles Unified School District, New York City Public Schools, Ohio, the Pittsburgh Public School District, Tennessee and the Tulsa Public Schools. States currently developing value-added models include Nebraska, Nevada and Pennsylvania.

Q: Will value added encourage teachers to only teach in schools with high test scores?

A: The value-added model accounts for students' prior achievement and other background characteristics. It therefore allows any teachers to be identified as high performers, even if their students start with low test scores. Value-added models adjust for prior test scores, and are designed to credit teachers not for high test scores in isolation but for the amount of progress their students made over the year, compared with the progress made by similar students. In this sense, value-added measures can "level the playing field" in ways that other measures that focus on static achievement levels cannot.

Indeed, a large body of research has shown that many teachers with high value added can be found in relatively low-performing schools. Recent evidence also suggests that such teachers bring their value added with them when they move to another school. For more information, see Mathematica's recent report on the Talent Transfer Initiative, which is the most rigorous study to date about the effects of moving effective teachers across schools: http://mathematica-mpr.com/Education/tti.asp.

Q: How will the proposed changes to the state tests affect value-added results?

A: A change to the state assessments will not affect the approach to estimating value-added results because a value-added model does not require that scores on the pre-test and post-test be directly comparable. Rather than directly comparing the scores on assessments from year to year, the value-added model estimates a student's typical-peer score on the post-test, based on pre-test scores and other information about the student.

Value-added and TLE

Q: When will this affect my evaluation?

A: Pilot value-added results will be provided to educators statewide during the 2013-2014 and 2014-2015 school years. These results will be used for informational purposes to help educators learn how to understand value-added measures and use them to drive improvement. Value-added results provided during the 2013-2014 school year will be based on instruction provided during the 2012-2013 school year and post-tests taken in 2013. The 2015-2016 school year is the first year that value-added results will be included as part of educators' full TLE evaluation rating. These results will be based on instruction provided during the 2014-2015 school year and post-tests taken in 2015.

Q: How are value-added results converted into a 1-5 scale for TLE? Is there a new conversion each year?

A: A teacher's value-added result will be converted to a TLE score between 1.0 and 5.0 based on a translation table. The table will be available for all educators to review as part of the technical documentation that is produced each year when value-added scores are released.

Q: Which teachers receive value-added results?

A: For the two-year pilot, teachers of the following courses will receive value-added results: math and reading in grades 4 through 8; algebra I, algebra II, geometry, and English III. The subjects and grades for which value-added results are calculated will be reevaluated after the pilot and are subject to change. Teachers must have at least 10 students who are eligible to be included in the value-added model to receive a value-added result. Eligible students must have both a valid post-test score from one of the grades and subjects listed above and a valid pre-test score in the related subject from the previous year. For grades 4 through 8 math, algebra I, algebra II, and geometry, the related pre-test subject is another math assessment. For grades 4 through 8 reading and English III, the related subject is another reading/English language arts assessment.

Q: How will teachers in grades and subjects that do not receive value-added results be evaluated?

A: Teachers in grades and subjects without value-added results will be evaluated using an alternate measure of student academic growth. For more information about student academic growth measures for non-tested grades and subjects, visit the TLE website.

Q: When will teachers receive their value-added results?

A: The first set of pilot value-added results, based on test data from the 2012-13 school year will be available in spring/summer 2014. In subsequent years, value-added results will be available to teachers in the winter of the school year after the tests they are based on. So for example, the value-added results based on 2013-14 test data will be available to teachers in the winter of the 2014-15 school year.

OSDE is building online dashboards to display all TLE data, including the value-added results. These dashboards will be piloted in fall 2014 and rolled out statewide beginning in the fall of 2015.

Capturing unique teaching situations, teacher and student transfers and absences

Q: How is student attendance, including ongoing course-specific absence, accounted for in roster verification?

A: The goal of roster verification is to allow teachers to review and (if needed) edit the roster data the school's student information system (SIS) has captured for the courses they teach. Although daily attendance is captured and reported through the school's SIS, roster verification offers teachers the opportunity to capture nuanced data about attendance that may not be accurately reflected otherwise. For example, if a student repeatedly misses a class due to related service provision and/or other ongoing conflicts, a teacher can use the roster verification process to make sure this repeated absence is reflected in the roster data that will be used to calculate value-added scores.

Q: Will students with IEPs be included in the value-added results for the regular classroom teacher or the resource teacher?

A: Students with IEPs who take the regular state assessments will be included in the value-added results for both the regular classroom teacher and the resource teacher. During roster verification both the general education teacher and the special education teacher will indicate their shared responsibility for delivering instruction to these students. The student scores will be weighted to reflect this shared responsibility when calculating the value-added results for both teachers. Students taking an alternate assessment will not be included in any set of value-added results. A teacher must have at least 10 students who are eligible to be included in the value-added model to receive value-added results. Consequently, resource teachers who do not meet this minimum will not receive value-added results.

Q: How do we account for co-teaching when determining a teacher's value-added?

A: The Oklahoma Value-Added model will address co-teaching using an approach called the Full Roster Method. This approach uses information about the instructional time individual students are with a teacher provided during the roster verification process. For example, when two or more teachers claim the same students during the same term at the same percentage, the Full Roster Method assigns each teacher equal credit for the shared students. Thus, solo-taught and co-taught students who are claimed by the teacher at the same percentage contribute equally to teachers' value-added estimates.

Q: Does roster verification start this year?

A: Districts are required to pilot roster verification in the spring of 2014 in at least a representative sampling of school sites. Although full participation is not required until spring of 2015, districts are strongly encouraged to participate fully in spring of 2014. The quality of teacher-student links, which play a central role in value-added analysis, depends largely on the roster verification process. Teachers who participate in roster verification will benefit by receiving more accurate value-added results sooner than teachers who do not participate.

Q: How does the value-added model address teachers in block scheduling classes?

A: The value-added result will reflect the amount of instructional time a teacher spends with each student during the year. Teachers will have an opportunity to confirm this information during roster verification.

Q: How do we account for extended teacher absence due to circumstances such as illness or maternity leave?

A: Through roster verification, the teachers themselves (or their administrator if the teacher is not present) will report any extended period in which they were not present for instruction. This also includes situations where the teacher is reassigned or changed positions. This data will be used to assign value-added results that accurately reflect the amount

of instructional time each student spent with a teacher. Districts may also address extended teacher absence through their local evaluation policies.

Q: If a teacher is reassigned to a different class within a school or moves to a new school or district during the school year, will they receive a value-added result?

A: This will depend on when the transfer happens during the year and whether or not they transfer to and from tested grades and subjects. Accurate student information system data and roster verification will help to ensure that teachers receive value-added results that correctly reflect their instructional time with students.

Q: What if we don't have a pre-test score for a student (for example, a student who transferred in from another state)?

A: Students without a valid pre-test score from the previous grade and year will be excluded from the calculation of a teacher's value-added result. To be included in the calculation, a student must have (1) a valid post-test score from a grade and subject for which value-added results are estimated and (2) a valid pre-test score in the related subject from the previous year. For math in grades 4 through 8, algebra I, algebra II, and geometry, the related pre-test subject is another math assessment. For reading in grades 4 through 8 and English III, the related pre-test subject is another reading/English language arts assessment. Students' pre-test scores in the related subject are typically the most important element used to estimate their typical-peer scores.