The Role of Value-Added Assessment in Michigan's Accountability System

Kwanghyun Lee, The Education Policy Center @ MSU
Bettie Landauer-Menchik, Policy Analyst,
The Education Policy Center @ MSU

Daniel Hu and Angela Sharp received excellent scores on the MEAP tests this year – scoring "Proficient" in both Math and Reading on the 7th grade tests. Daniel attends Farmington Hills Middle School and Angela lives in Flint. Clearly, they tested well compared to their classmates on the MEAP test that they took last spring, but do we know if they actually *learned* anything during their time in the Flint/ Farmington Hills school districts? Or would Angela and Daniel have done just as well on their MEAP tests regardless of which school they attended?

Michigan currently bases its accreditation system solely on the Michigan Educational Assessment Program (MEAP) tests, a "snapshot" measure which shows how a single student is performing at one point in time. But this snapshot doesn't reveal what Daniel or Angela really learned in school. It only tells how well they performed on a single test on a single day. In addition, MEAP scores are closely associated with family background. Family, neighborhood, and peer group characteristics affect student achievement cumulatively, over time. Because student learning is also cumulative, Daniel and Angela's progress cannot be measured only by a test they take in 4th and 7th grade. A single snapshot does not meet the needs of parents, teachers, and policy makers who want to assess how much progress students are making every year.

Michigan's policy makers should consider adopting a more accurate system of evaluating student progress by instituting a value-added assessment system. Value-added testing measures how much a student gains in specific content areas during one school year. Under a value-added assessment system, Daniel and Angela would be tested every year and their annual progress evaluated. Once Michigan adopts an accreditation system that includes annual testing, policy makers will be better able to distinguish between students who test well because of family and neighborhood effects and students who are progressing well because of excellent teachers and good educational policies.

How Can Value-Added Assessment Be Included in Michigan's Accreditation System?

With the Single Record Student Database (SRSD) and the Michigan Education Information System (MEIS), the state has established a framework for collecting information about individual students that can be used for value-added assessment. The components of a value-added assessment system that are still missing are (1) the implementation of annual testing, (2) the incorporation of annual testing data into the state's new accreditation model, and (3) the development of a system of

rewards and sanctions based on the learning that schools and teachers produce for Michigan students each year.

SRSD and MEIS

The SRSD and MEIS contain the necessary elements for an accreditation system that begins with the individual student. The SRSD contains student data, such as yearly achievement data, demographics, and attendance data. The MEIS contains teacher, fiscal, and performance data. By using both databases, data about what courses a student takes, teacher certification, and annual tests can all be linked to individual student records.

Implementation of annual testing

Annual testing can provide several important indicators for accountability: *the annual gain* in test scores from one year to the next, i.e. the "value-added" by a teacher and school; *the expected achievement of students controlling for prior achievement*; and *the mastery of standards*. Annual tests maximize accountability by isolating the effects of individual classrooms, schools, and districts. Once these effects are isolated, fair comparisons can be made between classes at specified grade levels, among buildings within a district, or between districts. Data can also be disaggregated by subgroups – by income, gender, or race/ethnicity. This allows districts to better identify students who are not progressing. Annual tests can provide districts with more up-to-date information on school performance than is currently available.

To implement annual testing, Michigan must decide which test to use. The state can purchase an "off the shelf " test, or it can construct its own tests based on the Michigan Core Curriculum Framework. If the state is interested in measuring how much students/schools improve test scores relative to national norms, then it should think about a *norm-based test* -- using either the national norm or a state norm. Examples of norm-based tests include the Stanford-9, the Iowa Test of Basic Skills, and the Metropolitan Achievement Test. Norm-referenced tests will make it easy to compare Michigan students with students in other states or across the nation, but they are not aligned with the Michigan Core Curriculum Framework.

On the other hand, if Michigan chooses to develop its own annual tests, they are more likely to be aligned with the standards and curriculum of the Michigan Core Curriculum Framework. This type of test is known as *criterion-based.* Michigan's MEAP tests are criterion-based with the "proficient" level set by the state every year. Other examples of criterion-based tests are the Texas Assessment of Academic Skills and North Carolina's End-of-Course and End-of-Grade tests. It would be possible for Michigan to develop a mini-MEAP test aligned with the Michigan Curriculum Framework. One disadvantage would be the time and expense necessary to develop a new system. However, a criterion-referenced test would provide far more useful information about how students are progressing towards state standards. An ideal testing program would provide data on both content and process mastery for each student.

Regardless of what testing program the state selects, the public needs to know what the state expects students to learn at each grade level. Then local districts can then design curricula, and teachers can prepare lessons that help all students excel on the state assessments.

Incorporation of annual testing data into the accreditation model

One of the most widely discussed models of value-added assessment was developed by William Sanders for use in Tennessee. Sanders' approach requires complex software to run a mixed model application to evaluate test scores and to measure teacher and school effectiveness. Recently, Sanders moved to the SAS institute in North Carolina as Director of SAS in School Inc. Michigan would have to contract with SAS to run Sanders' model. In order to use Sanders' model, the state would have to adopt one of the nationally norm-based tests, such as CTBS or SAT9. The advantage of this option is that Sanders' approach has already been tested elsewhere; the disadvantages are that Sanders has not shared his model for others to test, the findings are valid only for third to eighth grade growth, the model ignores initial achievement gaps, and the model itself is complex and difficult for people to understand.

The Dallas Independent School District (DISD) uses a standard regression model that analyzes each year's scores separately . The DISD model includes a number of different variables including measures of student background, in order to "filter out" the effects of school characteristics like the percentage of students eligible for free or reduced lunch program that predict student achievement.

Regression models are more familiar than Tennessee's mixed model, and the results are easier to interpret. The inclusion of student background variables in the DISD model address the concerns expressed by legislators and educators about fairness and equity. The DISD model focuses not on expected gains, but on expected achievement. The value added by the educator is no longer the amount of gain experienced, but the amount of discrepancy between the expected achievement and the actual achievement.

North Carolina also uses a regression model to determine a school's growth composite, but the North Carolina model includes fewer variables. The calculation of a school's expected growth composite is based on state norm from the 1992-93 and 1993-94 school years. The regression equation includes three factors:

a) the North Carolina average rate of growth in the respective grade and subject;
b) a statistical estimate of the "true proficiency" of the students in a school (i.e. Students who are more proficient might grow faster – that is how they got to be more proficient in the first place – they already grew faster); and
c) a correction for the movement of students' scores due to "regression to the mean."

The Texas model uses a simple algebraic growth measurement that includes income, race and ethnicity as variables. Texas evaluates schools and districts based on a common test-score target that must be met by each group. Each year, the targets are raised to motivate schools to continue to improve.

The state of Michigan can evaluate "value-added" models from these and other states, weighing the sophistication and accuracy of complex statistical models against the simplicity and transparency of models that can be understood by educators, parents, and legislators. Policymakers will have to balance the need for accuracy and fairness with the need for transparency.

Develop rewards and sanctions based on achievement and progress

An effective accreditation system needs to incorporate both "snapshot" measures of student achievement and measures of student learning gains. The system should encourage achievement growth for all students. At the same time,

the system should be fair to teachers by accounting for differences between the students they teach while at the same time encouraging them to close the gap between low and high achievers. Some schools that score low on MEAP may make exceptional progress with their students. Some schools that score high on MEAP may simply be taking advantage of students' prior learning, either in other schools or at home, and producing little in the way of learning gains. The incentive system must therefore include fair measures of achievement and of progress.

Schools that are identified as low-performing will require assistance either from the state, the ISD, or the local school district. Assistance might include support in the school improvement planning process or technical assistance. Schools that fail to improve after a specified period of time could have more severe penalties, including on-site review by state officials, increased technical assistance, or transfer and replacement of staff. Some states allow students in chronically low performing schools to transfer to other schools.

What Would Value-Added Testing Contribute to Michigan's Accountability System?

Value-added assessment enables teachers, schools, and districts to evaluate the impact that each year of instruction contributes to a student's progress. Daniel and Angela's gains can be compared against their scores in previous years. After calculating the gain in Angela's school in Flint and Daniel's school in Farmington Hills, the state can compare the gains with similar schools or with the state norm.

A value-added model can be relatively simple and easy to understand: Each student's progress can be compared to his/her own record over a period of years. Angela and Dan's achievement would be compared over years and across subjects – making it possible for the additive effects of teachers, schools, and districts to be analyzed. With a value-added system each student is compared to him/herself. Poor Michigan districts would not be able to use poverty as an excuse for students who don't progress, and wealthy Michigan districts would be scrutinized to ensure that lack of adequate progress is not hidden by already high scores .

Under a value-added assessment system, teachers, schools, and districts will have a greater

incentive to focus on teaching low achieving and average achieving students. Indicators, such as mean achievement level of schools, provide schools with distorted incentives to focus on high achieving students. With a value-added assessment system, the focus is likely to shift to average or below average-level students, since accountability will be based on the overall *gains* of students.

How can value-added assessment be adopted as part of Michigan's accountability system?

Other states that have already incorporated value-added assessment into their accountability systems can be critically evaluated as Michigan develops its own system. Extensive evaluation of the Texas, North Carolina, and Tennessee systems can help the state Board of Education choose among competing models.

The state's accountability system should include measures of improvement as well as achievement of high academic standards. Standards should be set that (a) challenge high-achieving students in all districts in Michigan, and (b) allow students from communities who enter school with fewer skills to advance. Annual testing will allow the state to measure and recognize improvement in low-income districts, at the same time recognizing districts where students continually meet high standards. The inclusion of valued-added measures can help to guard against "institutionalizing low expectations" for districts with high concentrations of low-income students.

Linking together high standards, value added assessment, the Michigan Curriculum Framework, and the Single Record Student Database, Michigan can establish an excellent foundation from which to build an accountability system. With such a system in place, the state and its citizens will be able to answer the question, "What did Angela and Daniel learn in school and who was responsible?"