

**Summary of Testimony at the Joint Committee Meeting
Senate Education Committee and House Education Committee
September 11, 2013
Deborah Loewenberg Ball**

My name is Deborah Loewenberg Ball. I am a former public elementary school teacher here in Michigan and currently professor and dean of the School of Education at the University of Michigan. I conduct research on mathematics teaching and learning. Every summer I also teach mathematics to fifth grade students, many of whom have been struggling with math in school. For the past 18 months, I had the honor of serving as chair of the Michigan Council for Educator Effectiveness (MCEE) as we worked to develop recommendations for a fair, transparent, and feasible educator evaluation system for Michigan's teachers and administrators.

My fellow council members and I took our charge seriously because we knew that the stakes were high for everyone—from teachers to school leaders to parents to the public. We examined research. We consulted with other states. We commissioned a pilot study. We engaged practicing educators. We requested expert guidance. We read and questioned and sought advice. Educator evaluation and student growth are complex issues and we appreciated the extra time we were given to produce responsible and defensible recommendations for Michigan's system.

My goal today is to provide you with an overview of the MCEE's recommendations and to explain the rationale for the key decisions we made along the way. I hope you will remember two things from my testimony:

First, the MCEE's recommendations are grounded in a deep belief in teachers—in the work that they do and in the crucial role that they play in student learning. In a growing body of research, study after study has shown the power of skillful teaching to affect student learning. Teachers who are prepared and supported to teach attentively are able to connect with students and make complex ideas and content learnable. They are able to form relationships with children and support them to master the academic and personal skills essential for success. Their professional skills enable them to care effectively for students' academic, social, emotional, and physical well-being and development.

Second, Michigan is poised to lead the way in implementing reforms to educator evaluation that put students first, that focus on improvement not punishment, and that contribute to the professional growth of all educators in the state. The recommendations that we made reflect these commitments and are grounded in a shared vision that my fellow council members and I established early on:

The Michigan Council for Educator Effectiveness will develop a fair, transparent, and feasible evaluation system for teachers and school administrators. The system will be based on rigorous standards of professional practice and of measurement. The goals of this system are to contribute to enhanced instruction, improve student achievement, and support ongoing professional learning.

This vision demonstrates our deep commitment to educators, to their continued growth and improvement as professionals, and to the academic outcomes and future success of Michigan's children. We believe that a system of educator evaluation should be grounded in improvement-focused feedback, not in punitive action. The system should, at its core, serve to raise the performance of **all** educators because this is what our students need and deserve. Our young people are the future citizens of our state and our nation; the quality of their education is our most important investment.

Let me turn now to summarizing briefly the MCEE's key recommendations.

Under the recommended system, both teachers and administrators would be evaluated on their practice and on their students' growth. For teachers, "practice" refers to the work that they do to prepare and

conduct instruction and to assess, communicate about, and improve students' learning. This includes working with students, administrators, caregivers, and colleagues. For administrators, "practice" refers to the work that they do to direct, operate, supervise, and administer the organizational and general education activities of a school or Local Education Agency (LEA). This includes setting curriculum, running buildings, evaluating teachers, directing action on their school improvement plan, and engaging students, parents, and teachers in the educational process. "Student growth" represents the change in students' knowledge and skill across time.

Under the recommended system, data from both practice and student growth would be combined equally—that is, 50% attributed to practice and 50% to student growth—to place teachers and administrators into one of three rating categories: professional, provisional, or ineffective. Professional educators are those who consistently exhibit the knowledge and capabilities expected of a skillful educator. Provisional educators are those who have uneven performance—who meet the standards in some areas but have other areas of weakness. This rating is intended to be a serious signal that the educator's practice requires significant improvement in specific areas. Ineffective educators have critical weaknesses and should be placed on urgent notice that improvement must be demonstrated within two years or they face termination. The goal of the system is to move an ever-increasing proportion of educators into the professional category through the use of feedback and targeted professional learning opportunities, thus ensuring that our state's children receive skillful instruction every year.

The MCEE opted for this three-category system instead of the four categories that were outlined in the statute for two reasons. First, we were concerned by the research we examined which indicated that the degree of measurement error involved would make it impossible to assign educators to four categories reliably. We were persuaded that trying to do so would make LEAs and the state vulnerable to legal challenges. Second, we were committed to orienting the evaluation system toward continuous improvement of educational practice. This commitment influenced the council to eliminate the "highly effective" category in favor of a model where a "professional rating" signifies teaching or leadership that is above the bar in terms of supporting students' learning, and yet is not "finished." We selected labels for the three categories that clearly separated practice that is unacceptable or in need of immediate improvement, on one hand, from practice that is meeting professional standards and that can be further refined, on the other.

A chart outlining the MCEE's recommended framework is on the following page.

For the practice portion, the MCEE recommended the use of an evaluative tool—an evidence-based "checklist" that guides evaluators to pay attention to the components of practice that actually matter for student learning. The use of such a tool would help to ensure that teachers' and administrators' practice is measured in a rigorous, professionally responsible, and legally defensible way.

Teachers' practice should be evaluated primarily through observations of their classroom instruction. But effective observation is more than just watching teaching and providing comments. Observation that is fine-tuned to key aspects of practice and that is systematic and not opinionated or idiosyncratic requires well-designed checklists and guides, as well as high-quality training for observers. The reliability and validity of inferences drawn from these observations—and, hence, their usefulness—depend on the rigor of both the tools and the process.

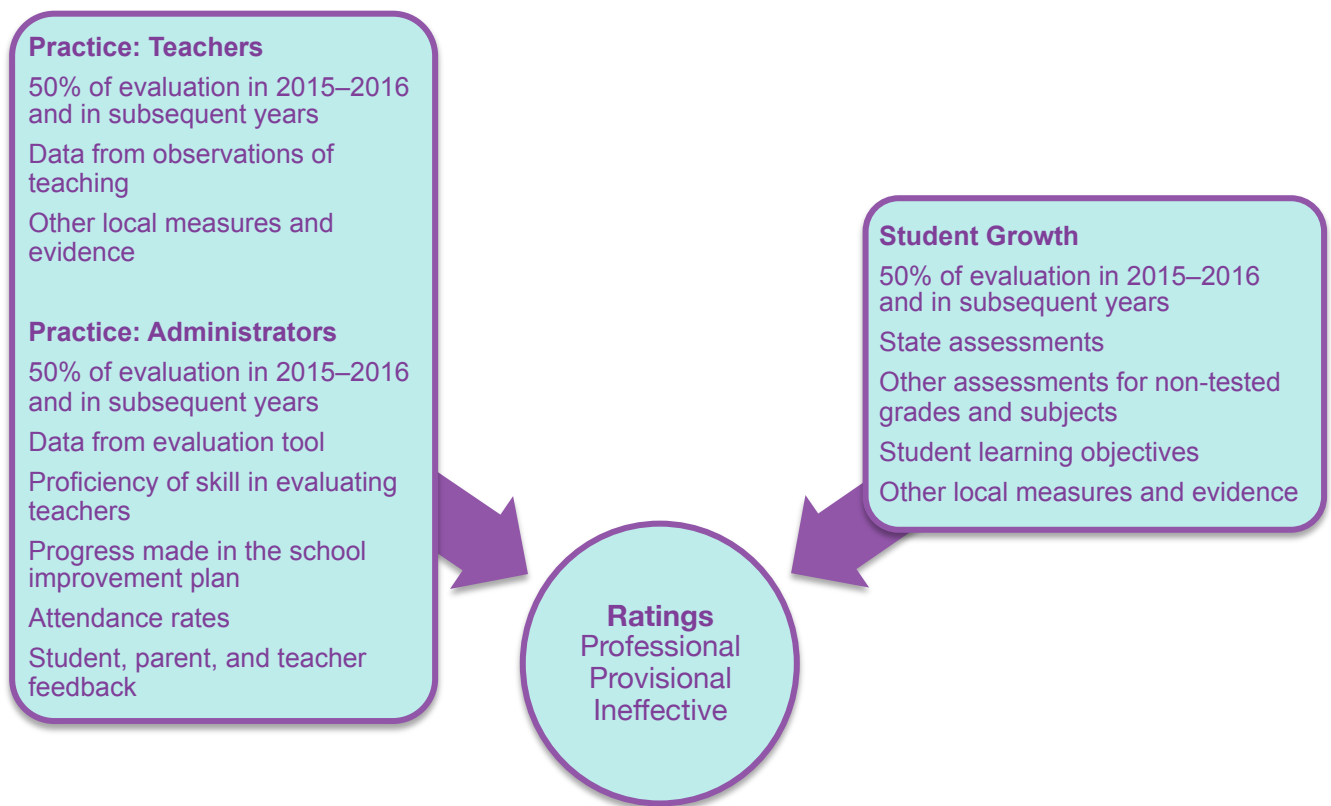
As a result, we recommended that the state select one of four observation tools piloted by the MCEE—Charlotte Danielson's Framework for Teaching, Marzano Teacher Evaluation Model, The Thoughtful Classroom, or 5 Dimensions of Teaching and Learning—to be the state tool, based on a competitive RFP process. Because current data from the pilot study suggest little significant difference between these tools, the MCEE opted to recommend an RFP process to put the state in the best economic position to solicit competitive bids from the four vendors and negotiate the best contract. It should also be noted that each of these tools has a strong research base and provides far better evidence about teaching effects than nearly all "home-grown" tools. We also recommended that the state provide sufficient base funding to support LEAs' use of the state-selected tool with full fidelity. This should include technical support and

vendor-provided training for all observers. If LEAs choose to use one of the other three piloted observation tools instead, they would pay for any expenses above the base funding supplied by the state for the state-selected tool.

Under the recommended system, well-trained observers, whether administrators or qualified peers, would use the tool to evaluate teachers multiple times across the school year. Evidence from the pilot study and from other research indicates that reliable estimates of a teacher’s performance depend on at least three observations. Each observation needs to be long enough to gather the information necessary to use the observation tool reliably and with fidelity, but does not need to be a full class period. At least one observation should be unscheduled.

For these observations to contribute to the continuous improvement of instruction, teachers must receive productive feedback oriented to the development of their practice. This feedback should be provided through conferences between the observer and the teacher and should reinforce specific strengths and highlight practices or skills that should be developed or improved.

Figure 1: MCEE’s recommended framework for educator evaluation



Observations are only one way to evaluate a teacher’s practice. Others include documents that support the observations, as well as other materials contributed by teachers, principals, students, or parents. Among the components used in other states are teacher self-assessments; professional development activities; educator growth plans (developed by teachers or administrators); structured reviews of student work; portfolios; and feedback from students, parents, and/or other teachers. Although many of these components could be effective in highlighting strengths and areas for improvement in teacher practice, the MCEE did not recommend their use be required and limited their optional use to no more than 20% of the practice section (or 10% of the overall evaluation).

In addition to requiring observations of teachers' practice, the recommended system would also evaluate teachers based on the growth their students demonstrate across a school year. This is perhaps the most complex and potentially controversial component of the MCEE's recommendations, so I will take some time to explain our thinking. The most important thing to bear in mind, though, is that we are talking about *growth*, not a particular level of attainment or achievement. All teachers should be producing growth in their students regardless of subject area or classroom composition. Student learning is the point of teaching.

The goal of the student growth portion of educator evaluation is to estimate the changes in a student's achievement that can be attributed to the instruction that student receives. One way to do this is to use changes in the student's scores on two administrations of a particular assessment, but this does not by itself provide a fair estimate of how much of the change is due to the teaching. Another approach is to provide evidence of the student's progress toward a set of articulated learning goals, such as is done with student learning objectives (SLOs) in non-tested subjects in a number of other states. Still another way to do this, when standardized tests are used to measure student learning, is to use statistical techniques called value-added models, or VAMs, that are designed to isolate the contribution of instruction by controlling for other factors that might impact students' growth (e.g., individual students' incoming achievement, race/ethnicity, socioeconomic status, gender, special education status, English language learner status).

There is, however, much scientific disagreement about VAMs, and some policymakers may question their use in high-stakes educator evaluation. We on the MCEE also recognized the well-documented limitations of these models, and cautioned against their overuse in educator evaluation. However, when comparing the use of VAM data to the alternative of locally developed data models of teaching effects, we determined that VAMs provide more reliable evidence. VAMs include many more statistical controls than what the vast majority of LEAs are able to build on their own, use more reliable assessment data, and involve more deliberate efforts to control for data quality issues that may bias the data. VAMs, while far from perfect, represent a better and fairer alternative to other means of ascribing test scores to educators. Therefore, we recommended that the state, in all cases that are possible and professionally responsible, produce value-added modeling scores for educators on state-provided assessments in the core content areas.

That being said, there are currently limited data available on which to calculate VAMs. In order to produce VAM scores, pre- and post-test data are needed. At this time, we have state-mandated tests only in grades 3 through 8 (MEAP) and grade 11 (MME). The MCEE's recommendations assume that the state will follow through on its plan to move all state-mandated standardized tests to the spring in the 2014–15 school year. Therefore the spring test in third grade will serve at the pre-test for fourth-grade teachers. Because there is no state-mandated test in second grade, there is no official pre-test for third grade, making fourth grade the first year VAMs can be calculated. And the only subjects that will initially be tested year to year during grades 3 through 8 are reading and math. In other words, for now, state-provided VAMs would only be available for teachers in grades 4 through 8 and only in reading and math. This will change, however, if the state moves forward with the Common Core and the aligned tests being developed by the Smarter Balanced Assessment Consortium.

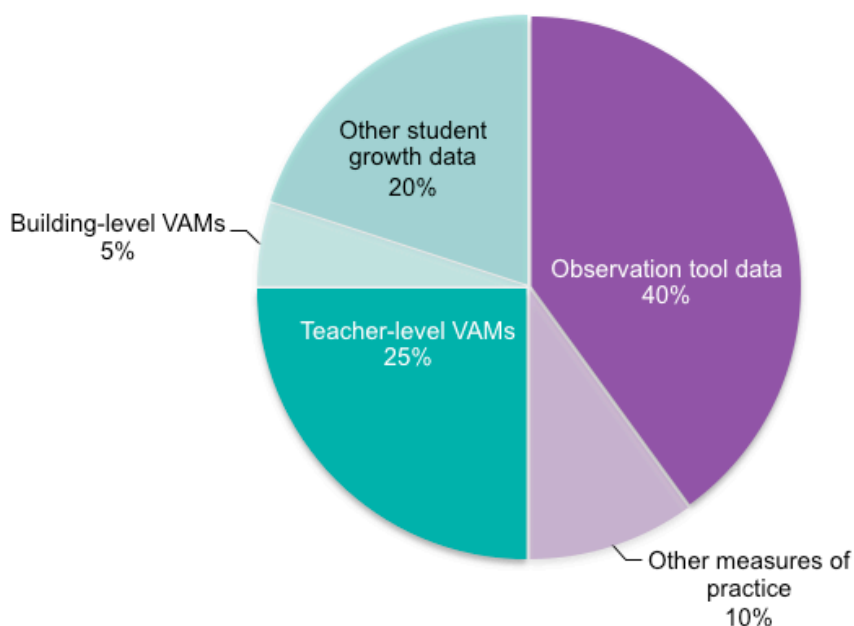
The Common Core and the Smarter Balanced assessments represent a much-needed step toward an educational system that would help all students develop the key skills and understanding needed for work and life in the twenty-first century. Having common goals for students' academic learning in core subjects will, for the first time, enable the development of high-quality materials and assessments to support skilled instruction. The standards will also enable much more effective preparation of teachers who can teach academic content effectively, as well as make possible a system of focused and fair evaluation of teaching and school leadership. This team approach to building excellent education will not—and could not—eliminate the many ways in which local contexts matter for students' development. What it can do is lift the quality of what we provide our young people and their teachers and stop the waste and weakness inherent in a “system” in which educational materials and practice are idiosyncratically invented in over 15,000 different ways without the benefit of the resources—research, expertise in the design and piloting

of curriculum tools and assessments—that are necessary for skilled professional practice. If medical practice were locally improvised in the way that education has too often been, the human costs would be terrible. It is time to understand that the costs to our young people of idiosyncratic practice have been similarly serious, and to advance the quality of education practice comparable with the progress we have made in medical practice in the last century.

Given the current testing regime, however, the MCEE recommended that for teachers in core content areas in grades for which there are growth data available from state-mandated assessments (again, currently reading and mathematics in grades 4–8, but likely to change over time), at least half of the teachers' student growth component should be based on state-provided VAM scores. In addition, the MCEE recommended that state-provided VAM or growth data in core content areas may be used in a teacher's evaluation using information from that teacher's students, even if the teacher does not teach in one of the core content areas. For example, a high school history teacher's evaluation may include her students' growth in English language arts because she is responsible for helping her students develop as readers and writers within her content area. Finally, in order to promote collective work on instruction within school buildings, we recommended the optional use of school-level VAMs for individual teachers' evaluations if there is a reasonable connection of the core content to the teacher's actual teaching assignment. However, school-level VAMs may not comprise more than 10% of the individual teacher's student growth component. In content area assignments for which there are no state-provided VAM or growth data available, the MCEE recommended that teachers be evaluated based on alternate measures of student growth that meet the guidelines for rigorous and appropriate measurement, such as vendor-created assessments, locally developed assessments, or SLOs.

One example representation of the proportional relationship of the types of data that can contribute to a teacher's overall rating beginning with full implementation of the recommended system in 2015–16 is below. This example applies to teachers in core content areas in grades for which there are growth data available from state-mandated assessments and assumes the minimum use of teacher-level VAMs (50% of the student growth section; 25% of the overall evaluation), the maximum use of building-level VAMs (10% of the student growth section; 5% of the overall evaluation), and the maximum use of other measures of practice (20% of the practice section; 10% of the overall evaluation).

Figure 2: Example of proportional data contributing to a teacher's evaluation



Under the recommended system, administrators would also be evaluated on their practice and on student growth. The MCEE recommended a similar RFP process to select either MASA’s School ADvance Administrator Evaluation Instrument or Reeves Leadership Performance Rubric to serve as the state tool for administrator evaluations. We again recommended that the state provide sufficient base funding to support LEAs’ use of the state-selected tool with full fidelity, with LEAs paying for any expenses above the base funding if they choose to use the other tool. In addition to the data generated by the evaluation tool, administrators would also be evaluated by at least the following evidence: their proficiency of skill in evaluating teachers; the progress made in their school improvement plan; attendance rates; and student, parent, and teacher feedback.

Student growth data for administrators could be drawn from several sources. The MCEE recommended that at least half the student growth portion of each administrator’s evaluation should come from building-level VAM scores where available (currently in reading and mathematics in grades 4–8, but likely to change over time). Both reading and math VAM scores should be weighted equally to make up the final building VAM score. The other portion of the student growth component would be determined at an LEA level.

For both teachers and administrators, practice ratings and student growth data would be combined to arrive at a final rating for each educator. One possible example of a set of evaluation ratings produced by considering both the quality of an educator’s practice and student growth is below.

Figure 3: Example of ratings produced by combining practice and student growth data

		Practice		
		Professional	Provisional	Ineffective
Student growth	Meets expectations	Professional	Professional	Provisional
	Does not fully meet expectations	Professional	Provisional	Ineffective
	Below expectations	Provisional	Ineffective	Ineffective

Although the MCEE worked diligently throughout our deliberations to recommend an evaluation system that ensures quality while also preserving some local control, we realized that some LEAs have used time and resources to develop their own evaluation systems. Under the recommended system, LEAs that are able to demonstrate the effectiveness of their local evaluation tools and policies would be welcome to apply for a waiver from the state system. However, because of the importance of fair and transparent evaluations for all educators, the MCEE recommended a waiver process that requires LEAs to submit research-based evidence of the validity and reliability of their locally developed or adapted systems. This waiver process should not be pro forma; it must be rigorous, structured, and able to withstand any legal scrutiny that might come from the high-stakes used of evaluative data.

Clearly, creating an educator evaluation system will require developing the relevant policies and processes and creating or adopting relevant instruments to be used. But compliance is not the point; skilled and principled use is. The system we have recommended will also require capacity-building—both across the state in LEAs as well as centrally—to understand, implement, and monitor the system, as well as infrastructure to house data, review materials, produce reports, and conduct program evaluation. A framework for monitoring and improving the system should be built in from the start. Moreover, we recommended that a complete audit of the system be conducted three years after full implementation. The audit should examine whether the system improves teaching and learning in the state and effectively supports ongoing educator learning and development. The audit should engage key stakeholders, including teachers, administrators, and parents.

There is no more important education reform than the one that Michigan is about to undertake. Every child in Michigan deserves skillful teaching, not just some of the time, but each and every year. And every teacher deserves the opportunity to develop and continue to refine his or her professional skill—to receive targeted feedback and professional learning opportunities to improve instruction. Making skillful teaching the norm will not only enhance the life chances and fulfillment of Michigan’s youth, but also boost our strength and capacity as a state.

On behalf of my fellow council members, I thank you for the opportunity to shape the future for Michigan’s schoolchildren. I look forward to working with you as you review our recommendations and use them to build Michigan’s educator evaluation system.