

GOOD FOR TEACHERS, GOOD FOR STUDENTS

The need for smart teacher evaluation in Michigan



The Education Trust—Midwest

TO THE POINT

- ▶ Smart teacher evaluation helps teachers understand their strengths and weaknesses and supports them in improving student learning.
- ▶ Despite their best efforts, many local school districts and charter schools don't have the resources or expertise needed to reliably evaluate teachers and give them the support and feedback they need to grow.
- ▶ By adopting state standards of evaluation, Michigan will help ensure that all models used in the state are reliable, technically-sound, and focused on improving teaching and learning.

Overview

As Michigan student achievement continues to fall behind a growing number of other states, it's clear that Michigan needs to support teachers better to improve instruction. Developmental feedback, in the form of a well-crafted, annual teacher evaluation, is an important first step toward that goal. Echoing their peers in other states, many Michigan educators say helpful, routine evaluations and useful professional development have been rare for much of their careers.

In an effort to give teachers the feedback and training they need to improve, the Michigan legislature passed a law in 2009 requiring local school districts and charter schools to evaluate all teachers every year, taking into account how much students learned. Since then, districts and charter schools have worked to develop their own evaluation models, often struggling mightily to ensure that the complexity and difficulty of teachers' work is taken into account.

Recognizing that struggle, the Michigan legislature returned to evaluation reform in 2011, creating the Michigan Council for Educator Effectiveness. The council of state-appointed education experts is charged with developing a statewide system of educator evaluation, including: Michigan's first common definition of what effective teaching looks like and a statewide evaluation model that any district or charter school in the state can use if it chooses. In addition, for those jurisdictions that want to develop their own models, the council is developing a set of state standards that all districts and charters would have to meet to have their models approved.

Still, a reasonable person might ask, is all this state-level action really necessary? Aren't our local districts and charter schools capable of deciding how to get their teachers the feedback and training they need to grow? To answer that question, the Education Trust–Midwest examined the teacher evaluation models now being used across Michigan.

The Education Trust–Midwest wanted to see how Michigan schools are responding to demands for developmental, technically sound — in other words, *smart* — evaluation practices that provide high-quality

professional development and feedback for our state's teachers. If we are to raise student achievement in our state, then we must do more to support and develop our educators' capacity to teach at higher levels. With this in mind, we reviewed local evaluation models adopted by 28 Michigan districts and charter schools of different sizes and capacities across the state. We then asked, "Do they measure up?" To help answer this question, we looked to best practices according to national research, lessons learned by other states and districts, and practitioners' recommendations. This report summarizes our conclusion, which is that most local models — despite the hard work that has gone into them — do not measure up to research-based standards for smart evaluation.

Included among our findings from the Michigan district and charter school models we examined:

- **Almost 20 percent** used checklist-style teacher observation tools with no opportunity for rich developmental feedback for teachers.
- **Almost half** allowed, or did not explicitly prevent, tenured or experienced teachers to go unobserved for an entire school year.
- **Only 18 percent** used the state's standardized tests to measure individual teachers' impact on student learning. State assessments are designed to measure how well students are learning Michigan's curriculum to ensure all students are getting rigorous, high-quality instruction to prepare them for an extraordinarily competitive global economy. Neglecting to use these available assessments leaves Michigan parents with no confidence that their children are learning what they are supposed to learn in school.
- **None** used a student growth or value-added model that was technically-sound enough to reliably gauge teachers' impact on student learning. Such measures are needed to provide rich feedback to teachers — and actually protect them from arbitrary evaluations.
- **The majority, 61 percent**, did not provide clear guidance to evaluators on how to combine the

“The process of evaluation has been yielding nice results for students — conversations are in more depth and it has forced us to have important conversations. We are grateful for this.”

— Scott Moore
Superintendent
Oscoda Area Schools

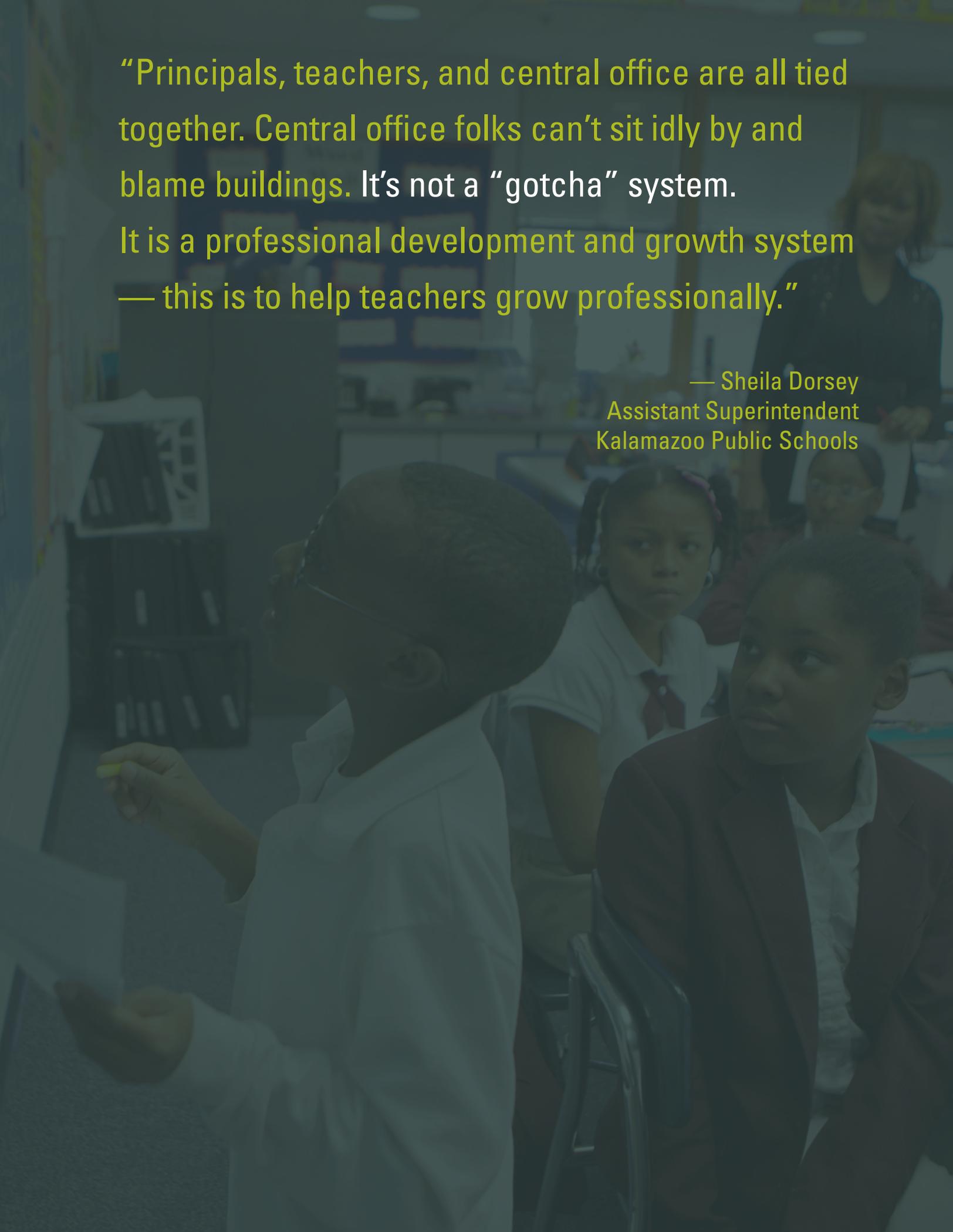
many measures of teaching performance into a final rating. This means administrators are more likely to produce unreliable or inaccurate final evaluation ratings — which may be risky for teachers, as these ratings will have a profound impact on their careers and futures.

- **No model** created a master or mentor teacher status or training to empower highly effective teachers to become observers in the evaluation process, which would help local schools manage the increased workload that meaningful evaluation may create.

Many of the district and charter school leaders we spoke with say they’ve long needed guidance to improve teacher evaluations. In other words, some state action is essential to protect everyone’s best interests — that is good for teachers, students and administrators.

To that end, the Michigan Council for Educator Effectiveness (MCEE) is developing a teacher evaluation system for the state. In the coming months, the Michigan legislature will have the opportunity to adopt the council’s recommendations. This report uses the lessons learned from local evaluation models already being implemented across the state to inform the work of the council and share these lessons with educators across the state, who often do this hard work in isolation. We also recommend standards the state should adopt to ensure that even those districts and charter schools that opt out of the state system meet minimum criteria for smart evaluation.

It’s also important to note that, despite the flaws found in these local systems, district and charter leaders across Michigan say the newfound emphasis on evaluation is helping teachers grow. Indeed, progress in other states and pockets of evidence from within Michigan have convinced us that it *is* possible to give teachers the kind of developmental, supportive feedback and data they need to truly excel.

A photograph of a classroom scene, dimmed and overlaid with text. In the foreground, a male teacher in a white shirt is pointing at a whiteboard with a yellow marker. Several students are visible in the background, including a girl in a white shirt and red tie, and a boy in a dark suit jacket. The text is in a bright yellow-green color.

“Principals, teachers, and central office are all tied together. Central office folks can’t sit idly by and blame buildings. It’s not a “gotcha” system. It is a professional development and growth system — this is to help teachers grow professionally.”

— Sheila Dorsey
Assistant Superintendent
Kalamazoo Public Schools

GOOD FOR TEACHERS, GOOD FOR STUDENTS

The need for smart teacher evaluation in Michigan

BY DREW JACOBS, SARAH LENHOFF, AND AMBER ARELLANO

INTRODUCTION: WHY DO WE NEED TO IMPROVE TEACHER EVALUATION?

For years, parents have known that teachers matter an awful lot when it comes to student learning. American society hasn't always acknowledged that. But in recent years, a growing body of research has shown how fundamentally important teachers are to students and to our state's and country's future. Indeed, when compared to every other in-school factor, teachers influence student learning more than anything else. More than class size, or even the school system itself, teachers have the most profound effect on how much students learn, and can even help close the achievement gaps that have persisted between groups of students.¹

Despite what we know about the power of teaching, national studies have found state and local school evaluation systems rate almost all teachers as "satisfactory." A recent study by the Education Trust–Midwest showed similar results in Michigan.² Traditional teacher evaluations have simply not provided the kind of constructive feedback that would help teachers improve.³ Not surprisingly, teachers nationally repeat the same refrain: Evaluation has not been helpful.⁴

INITIAL STEPS IN MICHIGAN

In 2009, Michigan began to take steps to remedy this problem. The legislature passed a law that required local school districts and charter schools to evaluate all teachers every year and take into account how much students learned as part of teachers' final evaluations. This put a new focus on student achievement and on supporting teacher development. However, unlike leading states, the law provided little guidance or state support on how to do this complex work well. Since then, districts and charter schools of varying size and capacity throughout Michigan have worked — often in isolation — to develop their own standards for good teaching and appropriate feedback, developing their own, unique ways of measuring teacher practice and student learning — and often struggling mightily along the way.

In 2011, recognizing an urgent need for assistance, the legislature passed pioneering tenure and teacher evaluation reforms to raise student learning and improve schools. This led to the creation of the Michigan Council for Educator Effectiveness (MCEE), a group of state-appointed education experts. Chaired by Dean Deborah Loewenberg Ball of the University of Michigan's School of Education, the council is charged with developing a statewide "opt-out" system of educator evaluation. That means that local school leaders can use the state evaluation model being developed by the MCEE, or they can seek approval for their own local models, which must meet new state standards the council is set to release in the spring of 2013. The council also is charged with developing

a new statewide value-added model for assessing student growth that all Michigan districts and charter schools will be required to use for a portion of their evaluations.

In November 2011, more than 300 districts and charter schools indicated their intention to opt out of the state evaluation model and develop their own instead. This report examines 28 of those models. (See sidebar and Appendix A to learn more about these 28 models and how they were selected for review).

HOW WE ANALYZED EACH LOCAL MODEL

To analyze the models, we looked to the latest, most widely accepted national research on teacher evaluation by policy and practice groups like TNTP (formerly The New Teacher Project), TAP: The System for Teacher and Student Advancement, and the National Council on Teacher Quality, and by respected education researchers such as Sanders and Horn; Goldhaber; and Rivkin, Hanushek, and Kain, among others.

Together, this research found that "smart" evaluations contained the following components:

- They are conducted annually
- Have clear, rigorous expectations that differentiate teacher performance
- Include multiple, structured, and comprehensive classroom observations
- Include sound measures of student learning growth
- Encourage constructive, clear, developmental feedback.

Because administrators and other evaluators are accustomed to rating almost all teachers as "satisfactory," smart evaluation models also give clear directions about how to combine multiple measures of performance into a final rating through what is known as a scoring framework. In Michigan's high-stakes educational environment, where evaluation ratings will dictate which teachers earn tenure and remain in the teaching profession, and which ones may eventually be dismissed if they do not improve after years of support, this work is critically important to the futures of thousands of professional educators and many more students.

Finally, research suggests that strong evaluation models include attention to both design and implementation. Smart evaluation takes time, expertise, and resources. Many educators across the state are worried about the new demands of quality educator evaluation, for good reason. One of the considerations in our study, therefore, has been whether these local evaluation

Drew Jacobs is a data and policy analyst, Sarah Lenhoff is assistant director of policy and research, and Amber Arellano is the founding executive director of The Education Trust–Midwest.

HOW WE SELECTED THE 28 DISTRICTS AND CHARTER SCHOOLS EXAMINED IN THIS STUDY:

The Education Trust-Midwest chose school districts and charters representing a range of public school experiences in Michigan's urban, rural and suburban areas. Our sample ranged in enrollment from 200 to 18,000 students. The districts and charter schools also varied in socioeconomic, racial, and special-education populations. Some — such as Oakridge Public Schools, Wayne-Westland Community School District, and Oscoda Area Schools — were previously cited as examples of best practices of teacher evaluation by the Michigan Department of Education and others.

In addition to analyzing the models themselves, we made multiple efforts to talk to the school leaders behind each model because we believe that context matters. Some of our most telling findings in this report come from these conversations. We are grateful for their courage, generosity and candor in speaking about this difficult work, the challenges they have faced, and how much they are learning.

Leaders in these 17 traditional public school districts and two charter school management companies agreed to talk with us: Bad Axe Public Schools; Berrien Springs Public Schools; Cadillac Area Public Schools; Choice Schools Associates; Dearborn City School District; Garden City Public Schools; Kalamazoo Public Schools; Kelloggsville Public School District through Kent ISD; Lapeer Community Schools; The Leona Group, LLC; Lincoln Park Public Schools; Livonia Public Schools; Oakridge Public Schools; Oscoda Area Schools; Rockford Public Schools; Romulus Community Schools; Rudyard Area Schools; Wayne-Westland Community School District; and West Iron County School District. We appreciate their contributions to our understanding of this important work, and we hope their insights will inform other districts and charter schools, as well as the ongoing work of the MCEE.

models help administrators and other staffers manage the potential increase in workload brought on by a new system.

Guided by this research, we assessed each evaluation model by asking five key questions:

1. Are teachers getting developmental, actionable feedback from their observations?
2. Is every teacher observed at least once a year by a trained evaluator? Are novices or low-performing teachers observed more often?
3. For teachers in grades and subjects with state standardized tests, are those test results used to generate technically sound estimates of a teacher's impact on student learning?
4. Do administrators and teachers know how to take the data generated from multiple measures of teaching — such as student growth, classroom observations, and student

surveys — and combine them to determine an accurate final evaluation rating? Is a clear, thoughtful framework provided to help them do this in a consistent way?

5. Are districts and charter schools finding effective ways to assist local educators with managing the increased requirements and workload, while administering consistently smart and reliable evaluations?

FINDINGS: THE STRUGGLE TO CARRY OUT SMART EVALUATION IN MICHIGAN

Question 1: Are teachers getting developmental, actionable feedback from their observations?

Finding 1: While some teachers are getting useful feedback from their observations, nearly 20 percent of the models we examined use checklist-style observation tools with no opportunity for the kind of rich feedback that will help teachers improve.

Classroom observation is the cornerstone of smart evaluation. Evaluators need to see teachers in action to know what they are doing well and how they might improve. But for observers to evaluate teacher practice accurately, they need a concrete vision of what good teaching looks like, with examples and descriptions of good practice in each of the four rating categories that Michigan now requires all schools to use: ineffective, minimally effective, effective, and highly effective. In other words, evaluators and school leaders need to clearly communicate what they mean when they say a teacher's practice is "effective" or "ineffective." As Sheila Dorsey, assistant superintendent at Kalamazoo Public Schools, told us: "Teachers want quick, actionable information."

In some places, such as the Dearborn City School District, the observation tools are both clear and descriptive, allowing teachers to receive rich, meaningful feedback that would help them improve. Dearborn administrators use an evaluation framework (or "rubric") that addresses 28 discrete teaching behaviors related to student learning that they can observe. As one example, administrators and teachers in Dearborn know what lesson clarity looks like at each of four performance levels, making it easier for administrators to conduct observations and to give constructive feedback, and making it easier for teachers to use that feedback as a guide for improving their practice (see Table 1, pg. 3).

In contrast, in White Cloud Public Schools, just north of Grand Rapids, administrators use an observation tool that gives no description of what each teacher rating category looks like, let alone guidance that teachers can use to improve their practice (see Table 2, pg. 3).

This kind of checklist-style observation protocol is typically of little use to classroom teachers because it provides a poor starting point for meaningful conversations about improving classroom performance. Smart observation tools give teachers precise and actionable information on current performance, what they can do to increase student learning, and a path for school leaders to support their teaching. They also allow schools to identify and use a teacher's strengths as an example

Table 1. Excerpt from the observation rubric that the Dearborn City School District submitted to the state in its bid to “opt out” of the state evaluation system.

DEARBORN CITY SCHOOL DISTRICT				
LESSON CLARITY	Ineffective	Minimally Effective	Effective	Highly Effective
	Lesson presentation has no defined structure nor is based on district curriculum.	Lesson presentation has some structure and is only partially based on district/state curriculum. Rarely is new information related to previous teaching.	Lesson presentation is clear, logical and based on district and state curriculum. Multiple techniques relate new information to previous teaching and accommodate student needs.	Model lessons are designed and shared with other teachers to improve understanding and practice across the curriculum.

Table 2. Excerpt from the observation rubric that White Cloud Public Schools submitted to the state to opt out of the state evaluation system.

WHITE CLOUD PUBLIC SCHOOLS				
TEACHER COMMUNICATES EFFECTIVELY.	Developing	Proficient	Accomplished	Not Demonstrated

for other teachers to follow. In Oakridge Public Schools in Muskegon, the district’s new approach to observations has already begun to improve conversations about practice. “We have taken a narrative approach to documenting observations which has been found to be more productive starting points for substantive dialogue about improving instructional practices,” said Superintendent Tom Livezey.

Sarah Earnest, Regional Human Resources Coordinator at Kent Intermediate School District, which helped Kelloggsville Public Schools near Grand Rapids design its model, said: “One of the biggest things I have learned is the power of the conversations that are had with teachers. That is where the learning occurs for the teacher.” A strong system of observation and feedback is one of the most effective ways to encourage powerful conversations between teachers and administrators.

The MCEE is piloting four promising, research-based observation tools in Michigan districts this school year. The results will help the council establish guidelines and recommend a statewide evaluation model that puts rich and meaningful feedback and collaboration at the forefront of teacher development. In the coming months, the legislature needs to step up and pass the council’s recommendations into law.

RECOMMENDATIONS FOR MICHIGAN

- The Michigan legislature should support new state standards requiring all district and charter school evaluation models to use observation tools that focus on the teaching behaviors most related to student learning, and contain clear descriptions of what teaching looks like within each level of performance.
- The MCEE must develop an observation tool that focuses

on the teaching behaviors most related to student learning in its statewide evaluation model. This tool must contain clear descriptions of practice at each performance level.

Question 2: Is every teacher observed at least once a year by a trained evaluator? Are novices or low-performing teachers observed more often?

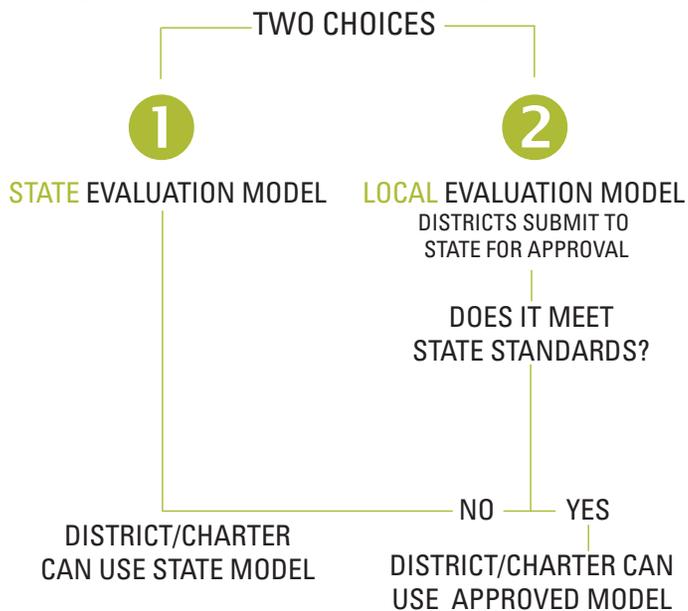
Finding 2: In roughly half the models we reviewed, teachers were observed at least once a year as part of an annual evaluation. However, about 46 percent of the models either allowed (or did not explicitly prevent) tenured or experienced teachers to go unobserved for an entire school year.

Teachers in their 20th year want to grow and improve as much as teachers in their first year. Teaching is demanding work and all teachers can strengthen some aspect of their practice. Even veteran teachers benefit from high-quality observations.

While about half of the models we analyzed required all teachers to be observed at least once a year, 46 percent either allowed some tenured teachers to go an entire school year without being observed or did not prevent this from happening. New teachers *should* receive more feedback than veteran instructors; most experts suggest multiple observations each year during the first several years of teaching.⁶ However, we shortchange our veterans if we don’t observe and share feedback with them at least annually.

The model used by Garden City Public Schools, west of Detroit, requires all teachers to have at least one annual observation and allows more observations for teachers who are new or struggling.⁷ Minimally effective or ineffective teachers in Garden City receive three or more observations annually.

MICHIGAN'S OPT-OUT EVALUATION SYSTEM



This graphic portrays the two teacher evaluation choices that will be available to local school districts and charter schools in Michigan. Systems can choose to use the state evaluation model, which is recommended by the MCEE, or they can opt to design their own model, which must meet state standards to win approval.

Alternatively, Madison District Public Schools in Oakland County has an evaluation plan for some tenured teachers that calls for no classroom observation at all. Teachers under this plan meet with administrators to review goals in September and again at the end of the year to discuss progress toward goals. But with no annual observations, these teachers do not get helpful feedback on their strengths and weaknesses that only classroom observation can provide. They don't receive an outside perspective on their teaching, which all professionals need to grow and learn. All teachers, no matter their tenure status or experience, should be observed at least once a year as part of a comprehensive evaluation.

RECOMMENDATIONS FOR MICHIGAN

- The MCEE must require annual observations for all teachers, both in the statewide model it's developing and in the standards it's crafting for districts and charters seeking approval for their own models.
- Struggling or less experienced teachers should receive at least three observations each year, though these may be of varying lengths and may be conducted by either administrators or specially trained expert peers.

Question 3: For teachers in grades and subjects with state standardized tests, are those test results used to generate technically sound estimates of a teacher's impact on student learning?

Finding 3: Only 18 percent of the models we examined use the state's applicable standardized tests to measure student learning. None of the models we examined used a technically sound growth or value-added model to gauge teachers' impact on student learning. In other words, the models we reviewed did not take into account students' performance when they enter the classroom or other factors that may impact student growth.

Two decades of research have demonstrated that teachers can have a significant impact on how much their students learn during a given year. Sanders and Rivers, for instance, found that students who achieve at the 50th percentile at age 8 will rise to the 90th percentile by age 11 if they have three consecutive years of high-performing teachers.⁸ The same students with three consecutive low-performing teachers would decline to the 37th percentile by age 11. That's a 53-point gap in achievement within three years.

But research also shows that even good classroom observers can't always distinguish the teachers who produce the most growth in student learning from those who produce little or no growth.⁹ That's why it is important for teacher evaluations to include both classroom observations and more direct measures of student growth whenever possible. Adding this second group of measures can actually protect teachers from arbitrary evaluations, because they provide objective information regarding a teacher's impact on student learning, based on how much a student learned during the school year while also taking into account other factors in a student's background, such as poverty. To do this, districts in Michigan must look to appropriate standardized tests and use a technically sound growth model, such as "value added," to isolate a teacher's impact on student learning.

Some of the local models we examined identified the best assessments currently available for measuring student learning, such as the Michigan Educational Assessment Program (MEAP) and the Michigan Merit Exam (MME). In Kalamazoo Public Schools, for instance, the evaluation tool spells out when state assessments should be incorporated into some teachers' evaluations: "Where a statewide assessment exists for the teacher's subject or grade level, that assessment shall be one of the measures used along with other comparable, rigorous measures approved by the principal in conjunction with the leadership team."

Conversely, in Oakridge Public Schools, teachers can be held accountable for student outcomes based on less useful or reliable factors, such as pass/fail rates. This is a less objective measure of student learning because it can be affected by factors beyond how much a student learned, such as attendance. Even in subjects and grades where state assessment data are not available, it's important for the MCEE to provide guidelines for measuring a teacher's impact on student learning. Instead of allowing teachers in one district to be evaluated on their students' pass/fail rates and teachers

in another district to be evaluated on student growth on approved assessments, the state must provide clear guidance that is consistent across Michigan. Without a uniform standard, teacher evaluations across the state will not be comparable; meaning that a teacher rated “highly effective” in Grand Rapids may only be rated “minimally effective” in Grand Blanc. If this is allowed to happen, parents will have no way of determining if their teachers are truly effective, or if their districts or charter schools simply set a low bar for performance.

For those grades and subjects in which students are tested by the state, the most widely accepted way to measure student growth is through “value-added” data. Value-added is a statistical measure that takes into account all of the student data available — including achievement scores from past years — to determine a teacher’s impact on student learning. These data look at the amount students grow during a year with a given teacher and compare that to how similar students grow elsewhere in the state. This is the most reliable way to measure the effectiveness of teachers in tested subjects and grade levels (typically about one third of all teachers) because it takes into account not only where students are at the end of the year but also where they started at the beginning. It also takes into account important factors that can affect student achievement, such as poverty. Why is this important? Accounting for these factors prevents teachers from being penalized for taking on challenging teaching assignments involving struggling or impoverished students.

Models like the one proposed by Berrien Springs Public Schools illustrate why value-added is so vital to making teacher evaluation reliable. Berrien Springs, in the state’s southwestern corner, doesn’t use a value-added model. Instead, it rates any teacher whose students, overall, perform “above the state average” as effective. In practice, this means that teachers who teach low-performing students will be vulnerable to misleadingly low evaluations, even if their students improve drastically during the school year. A sophisticated growth model such as value-added would credit those teachers who dramatically improve student growth, even when those students remain below the state average.

The MCEE is charged with developing a sophisticated value-added model for all districts and charter schools across the state. Even those seeking to win approval for their own evaluation model must use the state growth or value-added model at least in part to determine a teacher’s impact on student learning. This state model will boost the confidence of educators who, like Superintendent Tony Habra of Rudyard Area Schools in the Upper Peninsula, are “worried that teachers with a significant number of at-risk students on their caseload will end up being rated as less than effective” on their evaluations.

In addition to sound growth models, teachers — especially those teaching untested grades or subjects — would also benefit from other measures of student learning, such as uniform student learning objectives. These measures help to ensure that teachers are working toward standards-aligned goals that are reasonable and measurable. States such as Indiana, New York, and Rhode Island have adopted student learning objectives to gauge the performance of teachers who

“Principals are having great conversations with teachers that they never had before. They are spending more time in classrooms than they have in 10 years.”

— Jan Bermingham
Program Coordinator
Berrien Springs Public Schools

teach subjects that are not tested by state assessments. The MCEE faces a similar task — it must find a way to measure the impact of all teachers, even those teaching classes that are not tested by the state. The council could, for example, develop a list of student learning measures that districts and charter schools can choose from in addition to a growth model.

The district and charter leaders we talked with said they understand the importance of measuring student learning because it helps them determine what elements of teaching are the most important. “It used to be, if your class was quiet, you were doing a good job. Now, we look at whether students are learning,” said Cheryl Irving, assistant superintendent at Lincoln Park Public Schools, a district located just south of Detroit.

Every model in our sample fell short in soundly measuring student learning, however. Despite their hard work, we found no districts or charter schools used a value-added model to measure student learning for evaluations. Only about one-in-three models included *any* measure of student growth as 25 percent or more of a teacher’s evaluation.

RECOMMENDATIONS FOR MICHIGAN

- State leaders must ensure that all Michigan districts and charter schools adopt the state-provided growth model being developed by the MCEE and use it in their local evaluation models for teachers who teach tested subjects and grades, regardless of whether they use the state system or seek approval for their own. This is the only way to get comparative data on teacher performance across the state, which will help ensure all students — especially the lowest performing — have access to highly effective teachers.
- The MCEE must provide clear guidance about how to measure student learning for teachers of subjects and grades where standardized assessments are not available, the kind

Table 3. How Final Ratings for Rhode Island Teachers are Determined

This is one example of a tool administrators can use to determine teachers’ final evaluation ratings. In Rhode Island, administrators can take teachers’ professional practice scores from their observations and combine them with student learning scores to decide what rating to give teachers on their final evaluations.

		STUDENT LEARNING			
		4	3	2	1
PROFESSIONAL PRACTICE AND RESPONSIBILITIES	4	Highly Effective	Effective	Developing	Developing*
	3	Highly Effective	Effective	Developing	Developing
	2	Effective	Effective	Developing	Ineffective
	1	Developing*	Developing	Ineffective	Ineffective
		*Note. This disparity will trigger an immediate review. Source: Rhode Island Department of Elementary and Secondary Education http://www.ride.ri.gov/EducatorQuality/EducatorEvaluation/Docs/Teacher_Model_GB-Edition_II_FINAL.pdf			

of guidance that teachers of non-core subjects receive in states such as Indiana, New York, and Rhode Island. We recommend the state convene a group of leading educators, students and parents to develop guidelines or new assessments to measure student learning in subjects not tested by the state.

- The MCEE should also approve a second method of measuring student learning for teachers of both tested and non-tested subjects, such as a process for setting annual student learning objectives for districts and charter schools across Michigan.

Question 4: Do administrators and teachers know how to take data generated from multiple measures of teaching — such as student growth, classroom observations, and student surveys — and combine them to determine an accurate final evaluation rating? Is a clear, thoughtful framework provided to help them do this in a consistent way?

Finding 4: Of the models we examined, 61 percent did not clearly describe how to combine multiple measures of teaching performance into a final rating. This means administrators have to guess how to put together these new measures of performance, creating a risk that teachers’ final ratings will be inaccurate and that teachers will not receive the support they need to improve.

All measures of teaching practice — including observations, student growth, and measures like teacher and parent surveys — should be taken into account when determining teachers’ final ratings. But this is new work for administrators, who are used to basing evaluation ratings entirely on checklist-style observations and rating virtually every teacher “satisfactory.” Many administrators say they welcome guidance on how to combine these multiple, more nuanced measures to determine final ratings. A clear scoring framework would help them reliably determine teacher performance, so they can give teachers the feedback and support they need to improve.

But designing these scoring frameworks can be a struggle. “We are really wrestling with how to address the mismatch between student data and observation data,” said Sid Faucette of Choice Schools Associates, a charter school management company. Teachers are likewise struggling with these changes. “There is a tension for teachers about the difference between ‘highly effective’ and ‘effective,’” said Sarah Earnest, regional human resource coordinator from Kent Intermediate School District. “We were using ‘satisfactory’ and ‘unsatisfactory’ evaluations, and it was always, ‘I’m the best because I’m satisfactory.’ So we have to change some mindsets because all teachers are expecting to be ‘highly effective.’”

Rhode Island provides a good example of a strong scoring framework (see Table 3, above). Supported by both the Rhode Island teachers’ union and the Rhode Island Department of Education, the state’s scoring framework helps administrators combine different measures of teaching into a sound final rating. It also flags instances where there is a wide disparity between measures.

Scoring frameworks such as Rhode Island’s take the uncertainty out of evaluations and help administrators make reliable decisions about teacher performance. In those instances in which there are disturbing disparities in the scores — i.e., a teacher who is scored as “highly effective” in student growth, but is rated “ineffective” in the classroom observation — teachers can depend on external reviews to reconcile such discrepancies and ensure reliability and accuracy.

Sheila Dorsey, assistant superintendent of Kalamazoo Public Schools, said creating this type of external review could create more buy-in among teachers and administrators, both of whom are subject to performance evaluation in Michigan. “Administrators have to be randomly checked for their evaluations, so principals are trying to do a good job,” Dorsey said. “They really want to help teachers become better professionals. There’s a lot of coming together.”

But only nine of the 28 models we examined gave administrators clear guidance on properly assembling a final rating. For instance, the model developed by Island

City Academy, a charter school in Eaton Rapids, gives administrators little guidance on how to combine measures of teaching from observations, student learning, and parent communication into a final rating for each teacher. With no standards for scoring final ratings, administrators in the state are likely to fall back to the status quo: rating every teacher “effective” and not giving teachers the feedback they need to learn and improve.

RECOMMENDATIONS FOR MICHIGAN

- The MCEE must develop a meaningful scoring framework for the state evaluation model to provide a clear way for educators to combine each piece of the model into a final annual rating for each teacher.
- For districts that want to opt out of the state model, the legislature must establish a standard that requires every district and charter school to develop a clear scoring framework for administrators to determine final annual ratings for each teacher.

Question 5: Are districts and charter schools finding effective ways to assist local educators with managing the increased requirements and workload, while administering consistently smart and reliable teacher evaluations?

Finding 5: None of the local models we examined create a master or mentor teacher designation for highly effective teachers to become observers and assist with evaluations. Master or mentor teachers could help school principals with the workload of more frequent evaluation and provide more opportunities for feedback and collaboration with their colleagues.

Smart teacher evaluation is complex work that most districts and charter schools have either not fully embraced or mastered. It takes time and commitment to do it right, especially in places where annual teacher evaluations have never occurred. Sarah Earnest from Kent Intermediate School District reminded us that administrators now face vastly different expectations about their roles and how to measure success than they faced even a decade ago: “We’re changing the role of the building principal — they are the instructional leaders, and that needs to be at the forefront.”

“Principals are having great conversations with teachers that they never had before,” said Jan Bermingham, Program Coordinator at Berrien Springs Public Schools. “They are spending more time in classrooms than they have in 10 years.”

Most districts and charter schools in our sample are struggling with how to manage the increased demands of evaluation and the changing role of administrators, while still maintaining accuracy and reliability. Though some districts have devised ways to do this well.

In Lapeer Community Schools, for instance, the evaluation model makes clear that every teacher is going to be observed annually. Probationary teachers receive no less than two observations each year and tenured teachers are observed at least once annually. But Lapeer also gives administrators flexibility on the length and frequency of observations for effective, tenured teachers. This combination of specification

“It used to be, if your class was quiet, you were doing a good job. Now we look at whether students are learning.”

— Cheryl Irving
Assistant Superintendent
Lincoln Park Public Schools



and flexibility ensures that evaluations will still include necessary observation data but allows administrators to determine how much observation information is sufficient.

Other districts are also building flexibility into observations. Four observations of 15 minutes each, targeted on specific teaching behaviors, may actually be more do-able for some administrators than a single hour-long observation, while still providing rich information for evaluating teachers.

In Tennessee, the District of Columbia, and Hillsborough County, Florida, some schools have handled the increased workload of evaluation by allowing master or mentor teachers to observe and give feedback to their peers. Based in part on the widely admired TAP System for Teacher and Student Advancement, this method allows highly effective teachers to apply for master teacher status. Master teachers share both instructional and evaluation responsibilities with administrators. Not only does this model allow teachers more one-on-one time with evaluators, it also supports a path forward for highly effective teachers to share their expertise with others, increasing the school’s capacity and giving these teachers the recognition they deserve.

None of the districts or charters in our sample created a master or mentor teacher designation for highly effective teachers. But some of the models we studied indicated that school leaders are receptive to the idea. In Kelloggville Public Schools, for instance, the model refers to both administrators and “evaluators” conducting observations. And in Lapeer

Community Schools, the model allows for “teachers-in-charge” to evaluate teachers. These instances show that some districts and charter schools may be open to guidance from the state on how this new role for teachers might be fully embraced.

RECOMMENDATIONS FOR MICHIGAN

- The MCEE should recommend a process for identifying highly effective teachers as master or mentor teachers who can be trained to perform observations and give feedback to other teachers to help them improve their practice. Local district and charter schools that seek to use their own locally-developed models should provide for a master or mentor teacher option, too.
- The MCEE should recommend that all evaluation models start with tenured, highly effective teachers being observed at least once per year (more frequently for novice and/or struggling teachers), with the minimum number of observations increasing as administrators get more experienced with evaluation.

THE FUTURE OF TEACHER EVALUATION IN MICHIGAN

Most of the districts and charter schools in our sample worked hard to develop evaluation models that give teachers the data they need to improve their practice. And many of them are already beginning to see improvements in teacher morale and instruction. “The process of evaluation has been yielding nice results for students — conversations are in more depth and it has forced us to have important conversations. We are grateful for this,” said Scott Moore, superintendent of Oscoda Area Schools.

Still, building smart evaluation systems is challenging work. All of the 28 models we reviewed fell short in at least one way. This does not mean that smart teacher evaluation is impossible. It means that smart evaluation in Michigan will require the collective effort of state leaders, local educators, and educational experts. In leading states, state leadership has been instrumental in building systems that serve teachers and students well. In Tennessee, where schools have completed their first year under a new state educator evaluation system, students posted the biggest gains in learning that the state has seen under its current assessments. In Michigan, the districts and charter schools that are attempting this work say that their systems, whatever their flaws, are helping teachers grow. Indeed, progress in other states and pockets of evidence from within Michigan have convinced us that it is possible to give teachers the kind of supportive feedback and information they need to truly excel.

“Principals, teachers, and central office are all tied together,” said Sheila Dorsey from Kalamazoo Public Schools. “Central office folks can’t sit idly by and blame buildings. It’s not a ‘gotcha’ system. It is a professional development and growth system — this is to help teachers grow professionally.”

The Michigan legislature must finish the work it began in 2009 and pass the council’s recommendations into law, ensuring a comprehensive educator evaluation system that any

district or charter school can use, and a set of state standards that all local evaluation models must meet.

As the MCEE designs a statewide model for teacher evaluation, we hope it will learn from the experiences of educators highlighted in this report. By drawing on the strengths and weaknesses of local models already being used in Michigan, the council will be able to design state standards that all smart evaluation models must meet to improve teacher practice and, in turn, increase student learning. Good teaching is too important to our students’ future to leave this to chance. ■

Appendix A. This table documents how the district and charter school models in our sample measure up to research-based criteria for smart teacher evaluation. “Yes” indicates that the model met that standard.

District or Charter ¹¹	Is the observation more than a checklist, providing room for nuanced feedback?	Does plan require annual observation of all teachers?	If MEAP/MME data are available on a teacher’s class, does the plan require it to be used in the evaluation?	Is student growth measured in the 2011-12 evaluations? If so, what percentage?	Does plan give specific directions on how to score all components of the evaluations?
Bad Axe Public Schools	No	Not specifically	No	Yes, but extent is unclear.	No
Berrien Springs Public Schools	Yes	Yes	No	Yes, 24%	Yes
Cadillac Area Public Schools	Yes	Yes	No	No. 25% starting in 2012-13	No
Choice Schools Associates — 11 Charter Schools	Yes	Yes	Yes	Yes, 50%	No
Dearborn City School District	Yes	Yes	No	Yes, “significant portion”	Yes
Dearborn Heights School District #7	Yes	Yes	No	Yes, 38%	No
Garden City Public Schools	Yes	Yes	Yes	Yes, 25%	Yes
Island City Academy	No	Not specifically	Yes	Yes, 50%	No
Kalamazoo Public Schools	Yes	No	No	Yes, 25%	Yes
Kent ISD / Kelloggsville Public Schools	Yes	No	No	Yes, but extent is unclear.	No
Lapeer Community Schools	Yes	Yes	No	Yes, 8%	Yes
The Leona Group LLC ¹² — 12 Charter Schools	Yes	Not specifically	No	Yes, 50%	No
Lincoln Park Public Schools	Yes	No	No	No. 25% starting in 2012-13	No
Livonia Public Schools	Yes	No	No	Yes, based on the school improvement plan.	No
Macomb ISD ¹³	Unknown	Unknown	Unknown	Unknown	Unknown
Madison District Public Schools	Yes	No	No	Yes, 8%	No
Melvindale-North Allen Park Schools	Yes	Yes	No	No. 25% starting in 2012-13	Yes
National Heritage Academies — 43 Charter Schools	Yes	Yes	Unknown	Unknown	No
Oakridge Public Schools	Yes	Yes	No	Yes, but extent is unclear.	No
Oscoda Area Schools	Yes	Not specifically	Yes	Yes, 20%	Yes
Pontiac City School District	Unknown	Unknown	Unknown	Unknown	Unknown
Rockford Public Schools	Yes	Yes	No	Yes, 25%	Yes
Romulus Community Schools	No	No	No	Yes, 40% for tenured teachers, unclear for others.	No
Rudyard Area Schools	Yes	Yes	Unknown	Yes, 20%	No
Watervliet School District	No	Yes	No	No, 20% starting in 2012-13	No
Wayne-Westland Community School District	Yes	No	Yes	Yes, 30%	Yes
West Iron County Public Schools	Yes	Not specifically	No	Yes, 25%	No
White Cloud Public Schools	No	Not specifically	No	Yes, but extent is unclear.	No

NOTES

1. Robert Gordon, Thomas J. Kane, and Douglas O. Staiger, "Identifying Effective Teachers Using Performance on the Job" in *Path to Prosperity: Hamilton Project Ideas on Income Security, Education, and Taxes*, ed. Jason Furman and Jason Bordoff (Washington, D.C.: Brookings Institution Press, 2008), chapter 7. S. Paul Wright, Sandra P. Horn, and William L. Sanders, "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation," *Journal of Personnel Evaluation in Education* 11 (1997): 57-67.
2. Sarah Lenhoff, "Strengthening Michigan's Teacher Force: How a New Teacher Evaluation System will Better Equip Michigan Educators to Improve Student Achievement," (Washington, D.C.: The Education Trust, 2012).
3. Weisberg, D., Sexton, S., Mulhern, J., and Keeling, D., "The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness," (New York: The New Teacher Project, 2009).
4. Ibid.
5. TNTP, "Teacher Evaluation 2.0," (New York: The New Teacher Project, 2010). National Institute for Excellence in Teaching, "TAP Research Summary," (Santa Monica, CA: NIET, 2012). National Council on Teacher Quality, "State of the States: Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies," (Washington, D.C.: NCTQ, 2011). William L. Sanders and Sandra P. Horn, "Research Findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for Educational Evaluation and Research," *Journal of Personnel Evaluation in Education* 12(3) (1998): 247-256. Dan Goldhaber, "Teacher Pay Reforms: The Political Implications of Recent Research," (Washington, D.C.: Center for American Progress, 2009). Steven G. Rivkin, Eric A. Hanushek, and John F. Kain, "Teachers, Schools, and Academic Achievement," *Econometrica* 73(2) (2005), 417-458).
6. Thomas J. Kane and Douglas O. Staiger, "Gathering Feedback for Teaching: Combining High Quality Observations with Student Surveys and Achievement Gains," Measuring Effective Teaching Project, (Seattle, Wash.: Bill and Melinda Gates Foundation, 2012).
7. In August 2012, the Michigan Council for Educator Effectiveness announced the 14 districts that would pilot the new statewide teacher evaluation system in the 2012-2013 school year. Garden City Public Schools, a district in our sample, was one of the selected pilot districts. Because of this, the district's evaluation model will likely be different in the 2012-2013 school year than what is represented in this paper. The 13 other districts in the pilot are: Clare Public Schools, Leslie Public Schools, Marshall Public Schools, Mt. Morris Consolidated Schools, Montrose Community Schools, Port Huron Area School District, Big Rapids Public Schools, Farmington Public Schools, North Branch Area Schools, Cassopolis Public Schools, Gibraltar School District, Harper Creek Community Schools, and Lincoln Consolidated Schools. Leslie Public Schools and Marshall Public Schools also applied to opt out of the statewide system in November 2011. Please see the Michigan Council for Educator Effectiveness website for more details on the pilot: <http://www.mcede.org/resources/2012-2013-pilot>.
8. William L. Sanders and June C. Rivers, "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement," (Knoxville, Tenn.: University of Tennessee Value-Added Research and Assessment Center, 1996).
9. Thomas J. Kane and Douglas O. Staiger, "Gathering Feedback for Teaching: Combining High Quality Observations with Student Surveys and Achievement Gains," Measuring Effective Teaching Project, (Seattle, Wash.: Bill and Melinda Gates Foundation, 2012).
10. Tennessee Department of Education, "Teacher Evaluation in Tennessee: A Report on Year 1 Implementation," July 2012. http://www.tn.gov/education/doc/yr_1_tchr_eval_rpt.pdf
11. Note on school responses: Of the 28 districts and charters selected, 19 complied fully with our requests. National Heritage Academies was partially responsive. Madison (Oakland) Public Schools, Dearborn Heights, Island City Academy, Macomb ISD, Melvindale-North Allen Park, Pontiac, Watervliet and White Cloud did not respond to numerous requests for interviews or declined to talk with us.
12. The Leona Group LLC operated 22 schools in Michigan in the 2011-2012 school year, but only 12 of these schools applied to opt out of the state evaluation model.
13. The 21 Macomb ISD districts applying for exemption educate 128,745 students and include some of the largest districts in Michigan, including Utica (2nd), Chippewa Valley (7th) and Warren (8th).

ABOUT THE EDUCATION TRUST–MIDWEST

The Education Trust–Midwest works for the high academic achievement of all students at all levels, from pre-kindergarten through college. Our goal is to close the gaps in opportunity and achievement for all children, particularly those from low-income families or who are African American, Latino, or American Indian — in Michigan and beyond. As a statewide education policy and advocacy organization, we are focused first and foremost on doing what is right for Michigan students. The Education Trust–Midwest is affiliated with the national organization, The Education Trust, based in Washington, D.C. Ed Trust–Midwest is the second state office of The Education Trust.



The Education Trust–Midwest

306 S. WASHINGTON AVE., SUITE 400, ROYAL OAK, MI 48067
P 734-619-8008 F 734-619-8009 WWW.EDTRUSTMIDWEST.ORG