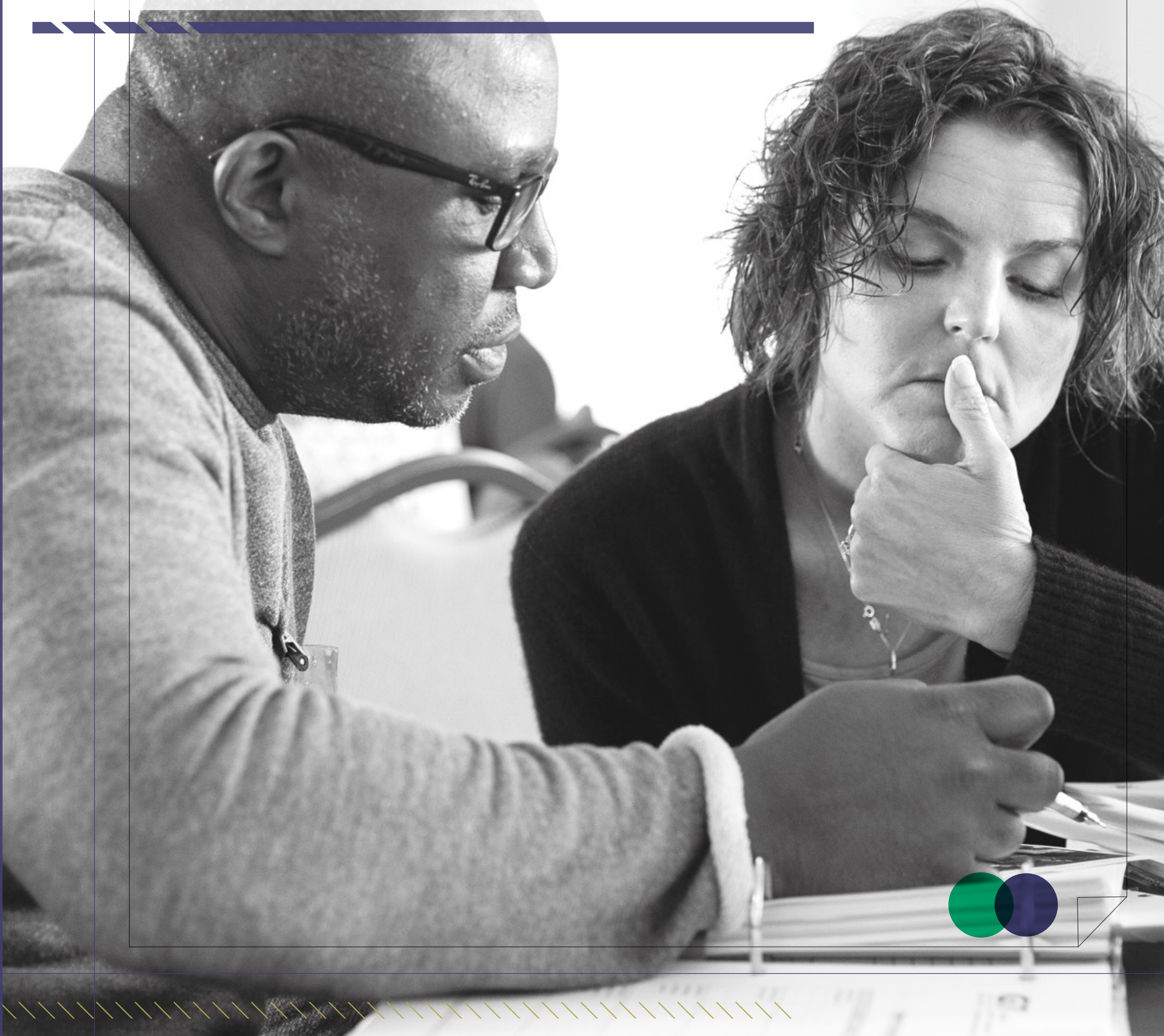
MATH IN COMMON®

Understanding Complex Instructional Change

Classroom Observations of Math in Common Districts



Suggested citation: Chu, H., Perry, R., Reade, F., & Marple, S. (2019). *Understanding complex instructional change: Classroom observations of Math in Common districts.* San Francisco, CA: WestEd.

Table of Contents

Introduction

Although mathematics standards have changed dramatically in recent years, teaching mathematics is as complex as it has always been. Some would argue that mathematics teaching has become even more complex, with the implementation of the Common Core State Standards for Mathematics (CCSS-M) (NGA Center & CCSSO, 2010), as teachers are being asked to make significant shifts in their instruction.[1]

Teachers report that they are incorporating math standards into their daily practice and are feeling positive about their efforts to do so (Reade, Perry, & Heredia, 2018; Perry, Marple, & Reade, 2017), but the education field still has little empirical documentation on exactly how math teachers are shifting their classroom instruction to align with the CCSS-M. Exactly what are math teachers doing in their classrooms to help students master the standards?

Part of the reason for the lack of data is the challenge of accurately measuring what happens during classroom instruction. The only real way to know what is happening in classrooms is through direct observation, and while it may be possible to get the gist of math classes through quick "drop-in" observations, it is ideal, for a systemic understanding of change, to use a valid and reliable observation instrument tied to specific elements of instruction. This sort of targeted instrument enables observers to obtain meaningful data and identify patterns in instruction across different lessons and teachers.

Regardless of who carries out these observations and analyzes the resulting data — teachers, principals, district staff, or partners from a research institution — it is challenging and time-consuming work. But this work is essential in order to gain knowledge of how the standards are being implemented in classrooms to support all students in achieving mastery of the CCSS-M. Without understanding of how teachers and students are responding to the standards, it is impossible to know what supports and course changes are still needed, from either a district perspective or a policy perspective.

Additionally, we frequently hear that there are not enough real-life examples of what the CCSS-M look like in classrooms when implemented well. Without examples of high-quality, standards-aligned instruction, it is difficult for educators to imagine how the standards should look and feel in their own classrooms, or to gauge their own progress. Carefully documented classroom observations can be a source of these sorts of real-world examples of standards-aligned instruction.

The Math in Common (MiC) initiative was launched to support CCSS-M implementation in grades K–8 in 10 California school districts. As part of its evaluation of MiC, WestEd conducted classroom observations in participating MiC districts to document K–8 teachers' instructional shifts related to the CCSS-M. The research staff visited elementary and middle school classrooms in nine MiC school districts, during the 2015–16, 2016–17, and 2017–18 academic years, to observe and analyze mathematics lessons, using an observation protocol adapted for this project. Participants from MiC teams

1 Recently, Lampert (2017) has contrasted ambitious teaching, focused on balancing conceptual understanding with procedural fluency and flexibility, with more traditional modes of instruction, focused on procedural repetition.

often joined us during the observations and debriefed with us afterwards.

Our preliminary learning from these classroom observation data was publicly reported in a blog post characterizing common structural features of highly rated lessons (Seago & Perry, 2017) and in case studies of incremental change in teacher practice over time (Seago & Carroll, 2018).

This report describes additional analyses of observation data on eight dimensions of classroom mathematics instruction. These analyses are drawn from our complete set of classroom observation data: 201 lesson observations, representing more than 130 hours of observation over three years. We begin the report by describing our classroom observation protocol and the dimensions of classroom instruction that we observed using this protocol. We then present our findings on the instructional variability that we saw across classrooms and districts. Next, drawing on classroom transcripts and observation data, we discuss what highly rated classrooms looked like across the various dimensions that we observed, and how administrators and others can support this sort of CCSS-M–aligned instruction. The report concludes with several recommendations for conducting effective classroom observations.

Our primary goal with this report is to share with teachers and administrators what we have learned about how particular elements of CCSS-M–aligned instruction look and feel when implemented effectively in the classroom. We also wish to stimulate discussion in the field about what kinds of information can best help educators understand standards implementation, and to share emerging insights from our experience trying to measure shifts in mathematics instruction.

WestEd's Protocol for Observing Classroom Instructional Shifts

The WestEd team wanted our classroom observations to contribute to the MiC initiative in several ways. First, in examining teachers' instructional practices related to the CCSS-M, we hoped to learn whether there were instructional patterns, within and across the participating districts, that could clarify teachers' progress toward implementing standards-aligned instruction. Most MiC district teams were themselves working on creating, piloting, and revising their own classroom observation protocols, and we hoped that, as they joined us on observations and discussed findings, we could support their efforts to collect their own valid, reliable, and useful classroom observation data. We also hoped that our formative analysis of classroom observation data could inform districts' teacher professional development, based on instructional successes and challenges revealed in the data.

To gather reliable and practical classroom observation data, we developed an observation protocol to measure several important dimensions of CCSS-M–aligned instruction. This section describes how we chose those dimensions, how we carried out the classroom observations, and what we learned about the interconnections among the dimensions of our observation protocol. Classroom observation is labor-intensive and complex work, and we hope that school district staff and others can learn from our experience.

Developing the observation protocol

In fall 2014, the evaluation team was certified on the Mathematical Quality of Instruction (MQI) instrument.[2] Although the MQI instrument was developed before the CCSS-M were released, MQI researchers found strong connections between the eight Standards for Mathematical Practice found in the CCSS-M and elements of MQI (Hill & Beisiegel, 2014).

The MQI instrument, designed for use on videos of classroom math instruction, encompasses 21 dimensions of teaching and learning, within four overarching domains. The WestEd team chose nine of the 21 MQI dimensions to pilot test in live (not video recorded) observations during spring 2015. We chose these dimensions because they

2 MQI was developed by researchers from the National Center for Teacher Effectiveness, the Center for Education Policy Research, and Harvard University. It was used as one of the primary observation instruments in the Measures of Effective Teaching study (Hill et al., 2008; Kane, McCaffrey, Miller, & Staiger, 2013). Getting certified on the MQI instrument involved many hours of participating in in-person professional learning and rating videos to calibrate our team's interpretations of the MQI codes with master coders.

seemed to be most closely aligned with the MiC districts' CCSS-M instructional goals.[3]

After piloting the observation protocol in 30 classrooms across five districts, the team selected five of the originally identified nine MQI dimensions to continue to use. (We found that four of the dimensions that we had originally selected were not producing clear or useful information about instruction.) In order to capture evidence of both teachers' and students' actions in the classroom, we slightly adapted the wording of three MQI dimensions to focus primarily on teacher work to support rich mathematics, while we used the other two selected dimensions to attend to students' mathematical thinking and behavior.

While we were developing our classroom observation protocol, several MiC districts were also beginning to use the Teaching for Robust Understanding (TRU) framework (Schoenfeld, 2014) for their own observations. After our pilot, we found that some important ideas about instruction that did not seem to be well captured through the MQI dimensions appeared in the TRU framework. Accordingly, in order to gather more robust evidence and to support our technical assistance to districts regarding their own observation programs, we incorporated three TRU dimensions into our protocol, for a total of eight dimensions — five MQI dimensions (see Table 1) and three TRU dimensions (see Table 2). These eight dimensions were ultimately chosen because they reflect three important instructional goals of the CCSS-M — student participation, mathematical practices, and conceptual understanding — and because they figured heavily in the work of the MiC districts.

Each dimension was rated on a rubric scale (a four-point scale for the MQI dimensions and a three-point scale for the TRU dimensions).[4] With this rubric, we hoped to gather useful observation data about how teachers were aligning their instruction with the CCSS-M, which would contribute to both our formative and summative evaluations of the MiC initiative.

Conducting observations in the field

Working in pairs, evaluation staff used our protocol and rubric to observe math lessons in grades K–8, and to gather evidence of the existence, and strength of implementation, of the eight dimensions in each lesson. WestEd staff paired up to observe 201 mathematics lessons between fall 2015 and spring 2018, often joined by members of each district's MiC team. The 201 lessons were conducted in 141 different elementary and middle school classrooms in nine MiC school districts. (Information about the observation sample is provided in Appendix B.) A district representative from the MiC leadership team selected the teachers to be observed each year, based on their grade level, availability, and interest. Each visit to a mathematics lesson lasted for approximately one hour (generally the full length of the lesson or mathematics period).

During the observation, the two WestEd observers took notes, which were then used to determine scores on each of the eight dimensions. After determining the scores individually, the pair met to come to consensus on final ratings for each dimension. For each lesson, they also produced detailed observation notes and summaries that

3 Acknowledging prior research on the complexities of implementing and observing classroom mathematics lessons (e.g., Lampert, 2001; Lewis & Tsuchida, 1999; Perry, Seago, Burr, Broek, & Finkelstein, 2015), we settled on observing nine MQI dimensions that fall into two larger categories: "richness of the mathematics" and "Common Core–aligned instruction." We recognized that we as observers would struggle to simultaneously observe more than nine dimensions during live classroom lessons.

4 MQI has a four-point scale of ratings: 1 = Not Present, 2 = Low, 3 = Mid, and 4 = High. TRU uses a three-point scale: 1 = Novice, 2 = Apprentice, and 3 = Expert. TRU subsequently eliminated the word descriptors from their observation tool, and also changed the name of one dimension from "agency, authority, and identity" to "agency, ownership, and identity." Appendix A includes more information on these rating levels.

Table 1. Mathematical Quality of Instruction Instrument Dimensions Included in the MiC Observation Rubric

Observation Category	Description
Linking representations	Teachers' and students' explicit, public (in small or whole groups) linking and connections between different representations of a mathematical idea or procedure.
Multiple solutions/procedures	Multiple solution methods occur or are discussed for a single problem. Multiple procedures for a given problem type occur or are discussed.
Mathematical sense-making	The teacher publicly attends to one or more of the following: the meaning of numbers, understanding relationships between numbers, connections between mathematical ideas or between ideas and representations, giving meaning to mathematical ideas, whether the modeling of and answers to problems make sense.
Student explanations	Students provide a mathematical explanation for an idea, procedure, or solution. Examples: Students explain why a procedure works; students explain what an answer means.
Student questioning and mathematical reasoning	Students engage in mathematical thinking that has features of important mathematical practices. There must be clear evidence of students engaging in such practices, such as: students provide counterclaims in response to a proposed mathematical statement or idea; students ask mathematically motivated questions requesting explanations (e.g., "Why does this rule work?"; "What happens if all the numbers are negative?").

Source: Adapted from Hill (2014).

Table 2. Teaching for Robust Understanding Framework Dimensions Included in the MiC Observation Rubric

Observation Category	Description
The mathematics	How accurate, coherent, and well justified is the mathematical content (including mathematical language)? Is there a clear mathematical goal for the lesson? How did mathematical ideas develop within the lesson for students?
Access to mathematics	To what extent does the teacher support access to the content of the lesson for all students? Who did and didn't participate in the mathematical work of the class, and how?
Agency, authority, and identity	To what extent are students the source of ideas and discussion of them? How are student contributions framed? What opportunities did students have to explain their own and respond to each other's mathematical ideas? How does the teacher respond to student ideas?

Source: Adapted from Schoenfeld & Teaching for Robust Understanding Project (2016).

described the activities and interactions of the lesson.

Observations were conducted twice in each academic year, in the spring and fall; some classrooms were visited repeatedly within a year or across years. (See the Math in Common evaluation report *Incremental Shifts in Classroom Practice: Supporting Implementation of the Common Core Standards—Mathematics* [Seago & Carroll, 2018] for case studies about changes observed in classrooms that were visited repeatedly across several years.)

Instructional Variability Across the Observed Classrooms

While our data set of observed lessons is relatively small and therefore is not representative of the full breadth of instruction happening in any single school or district, we believe that the findings in this report can still offer some useful snapshots of CCSS-M implementation as it played out in real classrooms. This data set helps us pose and answer some questions about instruction in the MiC districts: How much variation in lesson quality is apparent in our cross-district sample? How many lessons are being taught at an expert level in at least one of our measured dimensions?

The data from the 201 lessons that we observed illustrate variation in classroom instruction across observation periods, across the sample, and across districts. This variation is the result of differences in many factors, including instructional materials, grade levels, student assignments, class work structures (e.g., in small groups versus individual), selected tasks, student populations, and teacher experience. Given the relatively small sample of teachers, their varying interpretations of students' needs, and the many decisions that go into developing and implementing any given classroom lesson, we were not surprised to find such variability.

Tables 3 and 4 on page 8 show the frequency of ratings assigned for each of the eight dimensions across the sample of 201 lessons observed by WestEd staff. Several findings are evident from these data:

- The majority of lessons exhibited at least a moderate amount of student *mathematical sense-making* (134 lessons, or about 66 percent of the lessons, were rated Mid or higher). Of the five MQI dimensions, student *mathematical sense-making* received High ratings most frequently.
- Not Present (i.e., no evidence) was the most frequent rating for *linking representations* (68 lessons, or 34 percent) and for *multiple solutions/procedures* (75 lessons, or 37 percent).
- More than a quarter of the lessons (57, or 28 percent) received a High rating on *the mathematics*. Conversely, student *agency, authority, and identity* received the highest rating for only about one-sixth of the lesson sample (34 lessons, or 17 percent).
- More than half of the lessons were rated at the middle Apprentice level on each of the three TRU dimensions.
- The majority of lessons (108, or 54 percent) did not receive a High/Expert rating on any of the dimensions.

We also investigated the hypothesis that instruction might be more likely to shift toward higher ratings (i.e., more High or Expert ratings on the rubric) after teachers had had more time to learn about the standards and hone their practice. That is, we considered the possibility that rating

Table 3. Distribution of Lesson Ratings on the Five MQI Dimensions (*N* = 201)

MQI Dimension	Not Present	Low	Mid	High
Linking representations	68	45	56	32
Multiple solutions/procedures	75	39	67	20
Mathematical sense-making	23	44	76	58
Student explanations	41	55	61	44
Student questioning and reasoning	51	58	58	34

Table 4. Distribution of Lesson Ratings on the Three TRU Dimensions (*N* = 201)

TRU Dimension	Novice	Apprentice	Expert
The mathematics	31	113	57
Access to mathematics	20	132	49
Agency, authority, and identity	49	118	34

patterns would trend upward over the course of the observation periods. However, when we examined data for the full sample across all six observation periods, there were no statistically significant differences in lessons that were observed earlier in the initiative, versus those observed later in the initiative, on any of the dimensions.

These data indicating no improvements in ratings over the course of the initiative present a paradox. We know, from survey results and from district administrator reports, that teachers across MiC have made, and continue to make, significant efforts to shift their instruction to align with the new standards, but those efforts are not explicitly present in this data set. We think there are at least a couple of explanations for this.

One explanation is that, given our sample size, we saw only a very small representation of all of the types of mathematics instruction occurring across the district, or even occurring across the year in the classrooms that we did observe. Another explanation is that teachers do not make deep changes to their instruction overnight, but make these changes through a long series of powerful but incremental shifts. Teachers' subtle instructional shifts would not necessarily appear on a rubric that is only scaled with three or four points. However, as described in the 2018 MiC evaluation report (Seago & Carroll, 2018), these incremental shifts were observed, and richly described, when WestEd researchers visited the same set of classrooms for several years and developed relationships with the teachers in order to better understand the contexts and motivations for the changes they were making. We believe that, ultimately, our observation protocol was not sensitive enough to detect the real instructional changes that teachers were making.

Additional analysis of the observation data

We began our study knowing that reducing the rich, complex, and interrelated dynamics of classroom instruction to eight isolated dimension scores would present challenges. However, observations conducted with a reliable and valid instrument can reveal important information about instruction

Table 5. Observation Dimensions and How They Combined to Measure Common Components of Classroom Instruction

Observation Dimension		Statistically Identified Underlying Components
The mathematics	→	Access to challenging mathematics
Access to mathematics	→	
Student explanations	→	Student explanations that serve to support student agency
Agency, authority, and identity	→	
Mathematical sense-making	→	Sense-making and student questioning/reasoning
Student questioning and reasoning	→	
Multiple solutions/procedures	→	Multiple solutions/procedures
Linking representations	→	Linking representations

across classrooms, and we were curious to understand how the empirical data reflected classroom complexity.

To develop this understanding, we conducted two statistical analyses. We found that the eight dimensions were highly correlated with one another. Some of the dimensions were so closely linked statistically that they could be considered elements of the same phenomenon — the eight dimensions grouped statistically into five underlying components of classroom instruction.[5] Three pairs of dimensions were found to be associated with each other in a way that suggests that they address a common underlying component of classroom instruction. So, for instance, our analysis found that the dimensions of *student explanations* and *agency, authority, and identity* don't actually measure separate aspects of classroom instruction; they co-occurred so frequently in the scores that analysis revealed they actually both measure one underlying component of classroom instruction — *student explanations that serve to support student agency*. Two of the dimensions, *linking representations* and *multiple solutions*, were relatively independent of the others (see Table 5).

Although we did not initially intend for these statistical analyses to reduce the number of observation dimensions, we were intrigued by the results and by the implications for using these findings as the basis of a new protocol in the future. Reducing the number of dimensions could enable us to focus our observational efforts on collecting better evidence during observations, since it would reduce the number of simultaneous observation components in the complex atmosphere of a classroom. Further, if multiple dimensions are connected in classroom instructional practice, specific examples of such instruction could demonstrate ways to simultaneously improve multiple aspects of the instruction, and thus, the findings from the observations could help teachers and administrators get more "bang for their buck" by spotlighting multiple facets of instruction at the same time.

5 A principal components factor analysis was used to identify these five components. The results of this analysis are provided in Appendix C.

What Did Highly Rated Classroom Instruction Look Like?

This section describes the five components of classroom instruction that were derived from our statistical analysis (see Table 5) and why they are important to CCSS-M implementation. It draws from observed lessons that rated highly on each dimension, describing what the lesson looked like when implemented skillfully in a real classroom, and offering ideas for how districts can support each dimension. It begins by discussing the three pairs of connected dimensions, followed by descriptions of lessons that received high ratings on the two dimensions that our analysis found to be independent of other dimensions.

Connected dimensions

Access to challenging mathematics

This component of classroom instruction is measured by the following two connected dimensions in our observation protocol: *the mathematics* and *access to mathematics*.

How these connected dimensions support the CCSS-M

The mathematics dimension measures how well developed, correct, and grade appropriate the mathematics in a given lesson are, and whether the lesson offers a learning trajectory with a clear mathematics instructional goal that is grade appropriate. The *access to mathematics* dimension measures how broad and thoughtfully facilitated the access to these mathematics is, and which students are participating in a meaningful way in the intellectual work of the classroom.

The correlation of these two dimensions tells us that the teachers who are most skillful at rising to the mathematical content demands of the new standards are also likely to be thinking carefully about how to facilitate broad access to that content across their diverse classrooms.

Prior to the CCSS-M, U.S. classroom instruction typically followed a relatively standardized and procedural instructional model (Stigler & Hiebert, 1999; Nesmith, 2008). Students would listen to a lecture and then practice procedural steps, all typically with a focus on obtaining the right answer, rather than a focus on the path of reasoning to achieve that answer. This model of teaching and learning did not enable U.S. students to achieve as well in mathematics as students in other (especially East Asian) countries (although more recent long-term achievement trends are more promising for U.S. students; IEA, 2016a, 2016b). Research has further documented a long history of disparities in learning opportunities and outcomes in mathematics education, based on race, class, culture, language, and gender (e.g., O'Day & Smith, 2016; Reardon et al., 2018). In order to eliminate disparities and enable achievement gains, classroom instruction needs to create the conditions for all students to build and demonstrate mathematical understanding.

The CCSS-M assume that all students will have access to and opportunities to learn high-quality mathematics, regardless of their demographic characteristics, but the mathematics that students have access to — and, as a result, what can be observed in any classroom — result from decisions made by policymakers at the state, district, and school levels as well as decisions made by individual teachers. These decisions include decisions about which instructional materials to adopt at the district level; which classrooms students are assigned to; the sorts of additional instructional resources or support that are provided by school mathematics coaches; the sorts of knowledge and experience that teachers are supported to build; and how assignments are developed and implemented, based on the specific needs that teachers have identified for their students, in any given month or week of instruction.

What these connected dimensions look like in the classroom

In the 14 lessons that were highly rated for both *the mathematics* and *access to mathematics*, participation structures proved important. One of the most common features of evidence for this connected dimension was the variety of participation structures that facilitated broad student access to the intellectual work of the classroom. These structures included free-flowing small-group working formats, systematic or visibly random methods for calling on students in whole-class discussion, and norms for working in small groups.

Considerations for supporting these connected dimensions in other classrooms

- Lessons rated highly on these two dimensions tended to involve small

The View from the Field

Participation Structures Facilitating Student Access to Challenging Mathematics

Following are descriptions of two observed lessons in which teachers effectively promoted access to challenging mathematics by using specific participation structures with their students.

Student Peer-Engagement Structures

In a second-grade class, students were studying properties of polygons, including sides, angles, and vertices.

The class had many well-implemented structures to enable students to interact with one another and to enable the teacher to provide additional mathematical support. The variety of participation formats enabled students to access the mathematics in different ways:

- Students worked in pairs as they sorted pattern blocks and identified groups.
- After completing exercises in their workbooks, pairs of students went to the carpet without pencils and compared their answers, providing one another with oral feedback and, in some cases, corrections.
- Toward the end of the lesson, the teacher had students post their responses to a problem in a designated area by the door, and used these responses to select students to participate in a small-group discussion on the carpet.

continued on p. 12 >>

>> continued from p. 11

Structures for Whole-Class Discussion

In a third-grade class, students were working in small groups with problems that required them to distinguish area from perimeter.

The teacher waited until at least one-third to one-half of the class raised their hands before calling on a student. If not enough students raised their hands, the teacher had them discuss further in their group before asking the question again. A total of seven students then went to the front of the room and used a document camera to present to the whole class the shapes they had constructed on geoboards. They were able to point directly at their models as they shared their computational pathways to determine the area or perimeter.

groups of students whose individual work was scaffolded through clear roles and clear mathematical tasks that they accomplished together. Teachers and administrators should consider spending time early in the school year to plan out student roles for group tasks and set student norms for participating and speaking together, in order to ensure that teachers' classrooms support both meaningful mathematics learning and broad student access to that learning.

- Similarly, these lessons had classroom norms to promote wide participation in the mathematical work of the lesson. To broaden access and ownership for the work of the classroom, teachers can use strategies to call on a range of students to share their ideas with the class, such as having students randomly draw "equity sticks" or purposefully selecting students with interesting ideas or contributions, based on a quick formative check of students' individual or group work.
- In these lessons, it was clear that the teacher had established that mathematics was a collective project of discovery and not just a rote or procedurally focused practice. In the best cases (but not in all highly rated lessons), the teacher had defined a clear mathematical goal for the day's lesson, and that goal was clearly situated within the larger goals of the unit. District administrators can emphasize the importance of clear goals in lesson planning in order to strengthen the mathematical coherence of each lesson and unit.

Other observations about these lessons and standards implementation

- None of the 14 lessons that were highly rated for both *the mathematics* and *access to mathematics* were highly rated on the *agency, authority, and identity* dimension. Although teachers of these lessons promoted broad access to the mathematics, they did not explicitly attribute solution approaches to students (e.g., saying to the whole class, "Let's talk about Dante's method for solving this problem"). The practice of attributing solution methods to the students who use or describe them can distribute mathematical authority to students and foster student agency. In these lessons (as across the entire sample), the teachers were more frequently directing discussions or serving as the arbiter of what was mathematically correct, rather than positioning students as the intellectual leaders of the lesson.

Student explanations that serve to support student agency

This component of classroom instruction is measured by the following two connected dimensions in our observation protocol: *student explanations* and *agency, authority, and identity*.

How these connected dimensions support the CCSS-M

The *student explanations* dimension describes students' contributions toward offering a mathematical explanation for an idea, procedure, or solution, beyond their just recounting the steps they completed in working on a problem. The dimension also attends to students' expressing their thinking about *why* those steps work. The *agency, authority, and identity* dimension addresses the extent to which students are the source of ideas and of discussion of them. This dimension poses questions such as: Is it the teacher or the students who are the primary drivers of the conversation and the primary owner of the intellectual work of the classroom? How are student contributions framed during the lesson?

The CCSS-M state that one way for teachers to assess students' mathematical understanding "is to ask the student to justify, in a way that is appropriate to the student's mathematical maturity, why a particular mathematical statement is true or where a mathematical rule comes from" (NGA Center & CCSSO, 2010). In the CCSS-M, having students justify their answers is considered one of the key supports to build students' mathematical proficiency. Furthermore, rich student explanations about *why* an approach or procedure works allow teachers to gauge student understanding and make formative adjustments to instruction in the moment, such as covering a topic more deeply or moving more quickly, depending on how students' understanding is developing.

What these connected dimensions look like in the classroom

The four lessons that were rated at the highest possible level for both the *student explanations* and *agency, authority, and identity* dimensions had similar participation formats and task types. A common task structure (and a common professional development topic for MiC districts) was the "Math Talk." Generally, a Math Talk is an instructional routine in which students are given "purposefully crafted computation problems" to solve mentally (California Department of Education, 2013, p. 9). Solutions are shared and discussed by the whole class to support students' collective reasoning (see the sidebar *Math Talks in San Francisco Unified School District* on page 14 for an example). This whole-class sharing format can facilitate agency, as the teacher is guided to explicitly give students credit for their explanations. In some cases, students also respond to one another and evaluate how others' approaches were similar or different.

The portion of the second-grade lesson shown in the text box *The View from the Field: Students Explaining Multiple Strategies During a Math Talk* (page 15) comprised just nine minutes of instructional time, but it demonstrates not only students' sustained, public articulation of their own ideas, but also instances where students are reporting about, reflecting on, and evaluating the thinking processes of their peers. In these exchanges, the space given for students to elaborate on their explanations, combined with teacher moves to increase ownership of and attribution for ideas, facilitated the development of student agency and supported students' understanding of number composition and decomposition.

Considerations for supporting these connected dimensions in other classrooms

- Math Talks were a very common feature in lessons that rated highly on these connected dimensions. However, the strongest instructional features of the Math Talks often ended when the Math Talks did, and student explanations and agency, authority, and identity were not strongly present in other parts of observed lessons. Teachers and staff who design professional development should consider building on the success of Math Talks by extending rich student discourse and ownership of ideas into the remainder of the lesson, so that deep mathematical work is taking up more than just a few minutes of instructional time.

- When students are asked to explain their work, they frequently respond by recounting the procedure they carried out. Students providing justifications for their thinking, which we observed in the four most highly rated lessons, is a much more meaningful mathematical practice. Teachers should support students in emphasizing the "whys" of their approaches to solving math problems, by helping them think through and describe reasons that a particular approach worked.

MATH TALKS IN SAN FRANCISCO UNIFIED SCHOOL DISTRICT

1. Teacher presents the problem.

A problem is presented to the whole class or to a small group. Computation problems are always presented horizontally, so as to discourage fixation on the standard algorithm.

2. Students figure out the answer.

Students are given time (1–4 minutes) to silently and mentally figure out the answer. They signal quietly to the teacher (e.g., with a thumb up against their chest) when they have an answer. If they have found one way to solve in the time allotted, students are encouraged to think of another way.

3. Students share their answers.

A few students volunteer to share their answers, and the teacher records them, without judgment, on the board.

4. Students share their thinking.

With a partner and/or with the larger group, students share how they got their answers. The teacher records students' thinking and attaches their names to the solutions. As the students are sharing their thinking, the teacher asks questions that help them express themselves, understand one another, and clarify their thinking to make sense of the problem and its solution(s). Multiple ways of solving problems and the connections between them are emphasized.

Source: San Francisco Unified School District (n.d.).

Other observations about these lessons and standards implementation

- Among the four lessons that were rated at the highest level for both the *student explanations and agency, authority, and identity* dimensions, three were rated lower in terms of *access to mathematics,* because not all students were able to participate. While a Math Talk may enable sustained attention to student ideas in a whole-class format, not all students may have the chance to contribute or be engaged.

- In two of these lessons rated at the highest level for both the *student explanations* and *agency, authority, and identity* dimensions, ratings were lower in both *the mathematics* (Apprentice) and *student questioning and reasoning* (Low). In these two lessons, because the focus of discussions was on computational pathways, explicit attention was not given to larger mathematical ideas. These lessons also provided limited opportunities for students to engage in mathematical reasoning and questioning, which involve processes of conjecture and generalization that may require activity structures beyond those provided by Math Talks.

The View from the Field

Students Explaining Multiple Strategies During a Math Talk

In this example from a second-grade class that we observed, students shared multiple strategies that they used for computations, and compared and connected their approaches. This portion of the lesson contributed to its high ratings on both *student explanations* and *agency, authority, and identity.*

The Math Talk problem on the board was:

$$59 + 12$$

Students worked individually for three minutes, using mental arithmetic. The teacher then led the following whole-class discussion to have students describe their thinking and even explain each other's work.

Student 1: I used breaking each number into its place value.

Teacher: Explain how you did that.

Student 1: I decomposed the 59 into 50 and 9 and decomposed the 12 to 10 and 2. 50 + 10 = 60. 9 + 2 = 11. So I regrouped the 11. And the 60 becomes a 70. 70 + 1 = 71.

Student 2: I made a 10. I minused 1 from 12 and put it to the 9.

Teacher: Why?

Student 2: So I could make a 10.

Teacher: Did you think of 9 separately from 5? What does 5 mean?

Student 2: Six tens. I put the other 10 from the 12.

Teacher: Took the 1 from the 12 and put it there, so what does the 12 become?

Student 2: Eleven. Put 6 and 10 together = 7.

Teacher: So 6 + 10 is 7. I'll have a 6 and a 10, but I think you mean six tens. [Shows six tens and one ten plus 1.]

Teacher: I got a little confused with this part. When you told me 59 + 1 is 10 and a 6, I didn't know where 6 was coming from.

Student 2: Six coming from 9 and 1. Other ten was coming from the 11.

Teacher: What about this 6?

Student 2: I put together the 5 and the other 10.

Teacher: [writes "5 tens plus this one ten equals 6 tens"]
Any other strategies?

Student 3: I turned the 5; I decomposed the 59 into 9 and 5.

Teacher: Nine and 5?

Student 3: Nine and 50. And decomposed 12 into 10 and 2, and then I taked 2 and 9 and it equals 11, and then 10 + 50 = 60. 60 + 11 = 71.

Teacher: Any questions for [Student 3]?

continued on p. 16 >>

>> continued from p. 15

Student 4: He did the same thing as [Student 1].

Teacher: What was different between [Student 1] and [Student 3]?

Student 5: [Student 3] added 2 from 9.

Teacher: Over here, what did [Student 1] do different?

Student 5: Broke apart 11 into 10 and 1.

Sense-making and student questioning/reasoning

This component of classroom instruction is measured by the following two connected dimensions in our observation protocol: *mathematical sense-making* and *student questioning and reasoning*.

How these connected dimensions support the CCSS-M

The *mathematical sense-making* dimension captures teachers' work toward helping students situate the elements of mathematics — the numbers and procedures — in rich contexts and in awareness of their own production of knowledge. As described in the MQI instrument, *mathematical sense-making* involves focusing on concepts such as:

- The meaning of numbers
- Understanding relationships between numbers
- Relationships between contexts and the numbers or operations that represent them
- Connections between mathematical ideas or between ideas and representations
- Giving meaning to mathematical ideas
- Whether the modeling of and answers to problems make sense

Within the MQI, both teachers and students can be centrally engaged in these sense-making activities.

The *student questioning and reasoning* dimension measures how students are engaging in mathematical thinking connected to mathematical practices, such as wondering aloud why a rule works; offering theories, counterclaims, or conjectures in response to mathematical ideas; or forming conclusions based on observed patterns.

Some researchers have argued that teachers can strategically use diverse mathematical ideas introduced during class to support students' sense-making and conceptual understanding (Smith & Stein, 2011). Woodward and colleagues (2012), for example, recommend that teachers explain relevant concepts and notation in the context of a problem-solving activity and prompt students to describe how worked examples are solved using mathematically valid explanations. Supports for student mathematical sense-making can arise both from what a teacher allows or attends to during a lesson, and from students asking their own questions and sharing their reasoning publicly. In other words, evidence of sense-making and student reasoning can be built on the basis of what is publicly attended to, examined, compared, and contrasted during the course of a mathematics lesson.

What these connected dimensions look like in the classroom

Among the five lessons that were rated highly for both *mathematical sense-making* and *student questioning and reasoning*, contextualized story problems featured prominently.

These problems involved topics and concepts such as area/perimeter, modeling geometric patterns, and data for statistical analysis. In some instances of sense-making, observed classes focused on understanding the context of a story problem and interpreting it to clarify the mathematics of the problem. In other instances, sense-making was observed when students were interpreting the meanings of solutions. In addition to contextualized story problems, highly rated lessons on these connected dimensions also featured small-group formats, which provided students with additional opportunities to examine their mathematical thinking (e.g., questioning why an answer was not the one expected).

Considerations for supporting these connected dimensions in other classrooms

- Contextualized story problems can be the kind of rich task that provides multiple entry points for students with different abilities. Teachers can use them to assess students' current understandings of an idea, and they can provide signals for teachers about where to provide additional support for building students' understandings.

The View from the Field

Students Negotiating and Reasoning in Public Through a Contextualized Story Problem

In this example, students in a sixth-grade class were solving a story problem about selecting members for a bowling team. This portion of the lesson contributed to high ratings on both *sense-making* and *student questioning/reasoning*.

This example shows students applying their knowledge about measures of center (mean, median, and mode) and representations of data distributions to decide whom to add to the bowling team, based on the scores of three players over seven rounds of bowling.

As part of initially making sense of the problem, students were asked to pose questions, such as "If there was a round eight, would John score higher?" The subsequent whole-class discussion then considered the trends in the performance of various players across seven rounds of bowling:

Student 1: It can probably go higher; each round it got higher, except between 2 and 3.

Student 2: Every one except for the last one was in the 100s, so I am thinking it is going to go back.

Student 3: I disagree with [Student 1] and agree with [Student 2]. Based upon John's score, only three were rounds that go up.

Student 4: I agree more with [Student 1]. Because the odds of John getting a higher score are pretty slim. Six out of seven [rounds], he gets into 100s.

The students' sharing at the end of the lesson reflected and articulated different forms of reasoning, including computing measures of center, creating graphical representations, looking at ranges of scores in intervals, and deciding wins by round. There was a public discussion about rounding, and about its meaning when making decisions. Students wrote individually, talked in pairs or small groups, or shared/presented to the class their ideas about how to rank and decide across players.

- Teachers can also use other instructional strategies that emphasize the kinds of sense-making and questioning processes that we observed for these connected dimensions. A "three reads" protocol, for instance, is designed to "develop students' ability to make sense of problems by deconstructing the process of reading mathematical situations" (Fostering Math Practices, 2018). Teachers' sharing and recording of various students' understandings is a critical part of this routine. Similarly, international comparisons have illustrated how teachers' writing on black or white boards provides documentation of ideas that arise during a lesson, and thus may support students in summarizing their thinking and reflecting on what they are learning (Yoshida & Jackson, 2010; Stigler & Hiebert, 1999).
- Among the five most highly rated lessons, ratings for *student explanations* were also generally rated High (with only one lesson rated at Low). All five lessons, however, were rated at the Apprentice level for either *access to mathematics* or *agency, authority, and identity*. Some lessons included rich small-group discussion but not whole-group sharing out. We see small-group discussions as a fruitful

The View from the Field

Using Different Methods to Find Volume

In this example, students in a fifth-grade class explored geometric measurement and volume.

The class started with a number-talk task: "Using multiplication and addition, use different ways to express 360."

With emphasis on different methods established, the teacher launched students into the main task of exploring the layers of a series of six rectangular prisms, increasing in volume. Students were given the prompt, "Each layer of these rectangular prisms is 4 cubes by 2 cubes. How many cubes make up each prism?" Students were then asked to find the answer for each prism, and later to describe the prisms using symbolic notation (e.g., 1 layer = $4 \times 2 \times 1 = 8$ cubic units; 2 layers = $4 \times 2 \times 2 = 16$ cubic units). They were then asked to use this information to generate a formula for volume.

The teacher regularly checked students' understanding, asking them to specify what changed from one presented prism to the next. Students responded to these questions in different ways:

Student 1: The second figure has an additional layer.

Student 2: First it has 8 [cubes], so the next is multiplied by 2, then 3, and so on.

Student 3: Add 8 for each layer.

Student 4: I built figure F by adding 8 because it just adds another layer.

At the end of the lesson, the teacher summarized what one student said about how to find the formula for volume: "She looked at how many cubes were in one layer and multiplied by the number of layers [because] that's an easier way than counting cubes."

area for districts to build on — professional development can help teachers successfully scaffold rich small-group discussions and learn routines for surfacing the small groups' conclusions and questions to the rest of the class.

- Some lessons featured high-quality whole-group discussion with only a small number of students participating in the conversation, which limited the access of other students to those discussions. Despite the quality of the discussion, these lessons were not scored highly. Including more students in classroom conversations offers another important place for teachers to build on current success.

Individual dimensions

Lessons demonstrating high ratings for the *multiple solutions/procedures* and *linking representations* dimensions did not also necessarily demonstrate strong mathematical coherence or student agency; these individual dimensions were independent of the other identified dimensions. Ratings on these dimensions may have been heavily influenced by particular tasks chosen for the lesson. If teachers chose tasks that were more "closed" (i.e., having limited entry points) or that were scaffolded for students, there might have been less opportunity for either type of practice, and, thus, observers would have likely scored these two dimensions as Not Present. This section describes how each of these two dimensions supports the CCSS-M and provides evidence from highly rated lessons.

Multiple solutions/procedures

How this dimension supports the CCSS-M

The first CCSS-M Standard for Mathematical Practice requires instruction to engage multiple solutions to a problem. Understanding that there are multiple solutions, and being able to understand similarities and differences between different solutions, contributes to a rich view of mathematics that goes beyond the mechanical application of algorithms. Attending to multiple possible paths to solving a problem can contribute to students' mathematical proficiency. For instance, mathematically proficient students can often find various entry points into a problem, and can understand how others approach complex problems. The What Works Clearinghouse has found moderate evidence that exposing students to multiple approaches supports students' mathematical problem-solving skills and can lead to gains primarily focused in the area of procedural fluency (Woodward et al., 2012).

Our team adapted this MQI dimension to focus attention on the ways in which teachers acknowledged and used different student solutions during a lesson, consistent with current instructional guidelines such as the *5 Practices for Orchestrating Productive Mathematics Discussions* (Smith & Stein, 2011).

Linking representations

How this dimension supports the CCSS-M

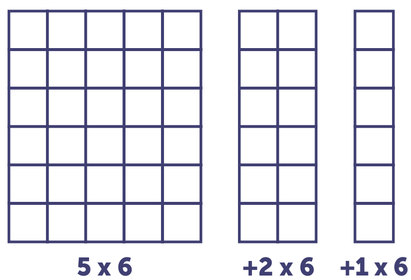
There is strong evidence that teaching students to use visual representations supports their mathematical problem-solving skills in grades 4–8 (Woodward et al., 2012). Instructional materials and standardized assessments often encourage students to show their work in multiple modes (e.g., in words, numbers, and pictures). The intention is for students to build flexibility in their problem solving, so that they are able to use different types of representations to explain and justify their understanding of a problem or how they arrived at an answer (Pape & Tchoshanov, 2001). But the reality is that when instructional materials prompt students to draw "pictures," it is possible for students' representations to display little or no corresponding mathematical content (e.g., in a perimeter problem about a cat walking along a fence, to encourage students to think about

The View from the Field

Linking Arrays, Mathematical Expressions, and Verbal Descriptions of a Math Problem

This third-grade lesson, which scored highly on linking representations, was focused on students using the distributive property and arrays to solve multiplication problems. Prior to the main lesson task, the teacher engaged students in a brief number talk: "Looking at an array, how many ways can we count groups of dots?" Six students described how they grouped the dots of a 4 × 3 array, and expressed their idea using a multiplication expression (e.g., "2 groups of 2 × 3" or "(2 × 3) + (2 × 3)"). As students shared, the teacher drew a diagram and recorded the number sentence, then verbally connected the numbers in the sentence to the grouping in the drawing.

For the main lesson task, students were given four arrays and asked to cut each array into pieces and use the distributive property to determine the mathematical expression. Many started with 6 × 8 arrays. In her discussion with the class, the teacher used student work that showed an array cut into 3 pieces and labeled as follows:



The class discussion focused on how this drawing was linked to the original 6 × 8 array, the expression (5 × 6) + (2 × 6) + (1 × 6), and the verbal description of the array (e.g., 6 groups of 8).

the distance *around*, students might draw a detailed picture of the cat, instead of focusing on the fence dimensions that are needed to compute perimeter).

Our team adapted this MQI dimension to focus attention on the ways in which teachers specifically drew connections among representations, to point out the common mathematical features of the representations and to support students in noticing these features and making similar connections themselves.

Conclusion and Recommendations

While classroom observations offer great potential for catalyzing improvement across district systems, various obstacles make them challenging to conduct and draw useful information from. Teachers often (correctly) perceive classroom observations as related to accountability rather than collective learning, which decreases the likelihood that they will open their classrooms. In addition, union contracts may place limitations on who can be in teachers' classrooms and for how long. There may also be insufficient staff with both training and authority (e.g., school principals) to conduct observations, or observation time cannot easily be spread across the large number of individuals who could benefit from formatively oriented observation.

Other concerns involve what is done with the information gained from observations and how useful that information can really be — who should provide feedback on instruction, how they should provide that feedback, and on which elements of instruction they should focus. There are also the challenges of developing or finding a reliable and valid instrument to measure the instructional elements of interest, and of using the resulting evidence to inform decisions about professional development and other investments in instructional change.

Given these challenges in carrying out effective classroom observations, it is no wonder that classroom-observation evidence on the success of standards implementation at the classroom level is not readily available. We have data on student outcomes, but we still know little about what teachers do in the classroom that influences those outcomes.

Our observations provide only a partial, yet still enlightening, view of how the standards are being implemented in classrooms. For example, we observed a fair amount of student mathematical sense-making, which is a key goal of CCSS-M teaching and learning. However, the connected pair of dimensions that comprised the *access to challenging mathematics* component were both rated highly in only 14 of the observed lessons (7 percent of all observed lessons), and the two dimensions that comprised the *student explanations that serve to support student agency* component were both rated highly in only 4 observed lessons (2 percent). Both of these connected pairs of dimensions reflect top-level goals of the CCSS-M, and the relative rarity with which we observed strong examples of these dimensions stands out as a call for further action.

To this end, we offer several recommendations to the field, and particularly to district practitioners, for conducting observations relevant to improving instruction:

- **To document school and district implementation progress on a larger scale, leverage existing relationships.** Our conclusions are limited by our small sample, but district staff may be better positioned than our external team was to visit more classrooms on a more frequent basis, in order to build a stronger body of evidence of instructional shifts and standards implementation.

- **Build greater classroom openness to facilitate broader evidence-gathering about instruction.** Some districts may have a legacy of observations being used for performance evaluations in a way that eroded trust with teachers. Districts should invest in repositioning observations as a way for adults to work together to support learning and improvement in a collegial, rather than punitive, sense.
- **Specify a purpose for an observational rubric and narrow the rubric down to a manageable set of dimensions of classroom instruction to observe.** Teaching and learning is complex; every dimension of a teacher's approach to mathematics, or of a district's mathematics vision, cannot be observed and measured simultaneously. Narrowing the focus of the observation, in order to generate useful data to guide improvement efforts, is crucial. In addition to being more efficient, focusing on just a few rich dimensions of teaching and learning — in both professional development and observations — may better reflect the coherence and complexity called for by the standards. The five components of classroom instruction described in this report, which were derived from existing frameworks, may be one starting point for districts to begin to understand instructional ideas as they connect in practice.
- **Build systems to enable district and school staff to observe and reflect on classroom observation data, and to act on those data accordingly.** Observation data are among the most actionable data that schools and districts can hope for. For example, our data revealed that shifting from having teachers explain math problems to having teachers support students in explaining and justifying their mathematical thinking may be one of the more difficult instructional shifts required by the standards. Other data can suggest similar target areas around which districts can organize professional development.

There is much more to be learned, within and across California districts, about how teachers are implementing the CCSS-M in dynamic classroom environments. Collectively, those of us in classrooms, district offices, and research institutions who are concerned about standards implementation have our work cut out for us, not only in building better systems of observation, but in using the data from those systems to better support instructional improvement. We hope that evidence from our study will enable other educators and administrators to prioritize this type of classroom observation–based evidence gathering in order to learn more about and improve standards implementation.

References

California Department of Education. (2013). Instructional strategies. *California Mathematics Framework*. Sacramento, CA: Author. Retrieved from https://www.cde.ca.gov/ci/ma/cf/documents/mathfwinstructstrategies.pdf

Fostering Math Practices. (2018). *3 reads (MP1)*. Retrieved from http://www.fosteringmathpractices.com/3-reads/

Hill, H. C. (2014). *Mathematical Quality of Instruction (4 point version)*. Ann Arbor, MI: Learning Mathematics for Teaching, University of Michigan.

Hill, H., & Beisiegel, M. (2014, May). *Mathematical Quality of Instruction (MQI) suite of tools*. Presentation given to Math in Common participants, San Francisco, CA.

Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An exploratory study. *Cognition and Instruction, 26*, 430–511.

International Association for the Evaluation of Educational Achievement (IEA). (2016a). *As global study TIMSS turns 20, new results show East Asian students continue to outperform peers in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center at Boston College. Retrieved from http://timss2015.org/wp-content/uploads/2016/T15-Press-Release-FINAL-11-29.pdf

International Association for the Evaluation of Educational Achievement (IEA). (2016b). *Highlights from TIMSS 2015* [Video]. Chestnut Hill, MA: TIMSS & PIRLS International Study Center at Boston College. Retrieved from http://timss2015.org/#/?playlistId=0&videoId=0

Kane, T., McCaffrey, D., Miller, T., & Staiger, D. (2013). *Have we identified effective teachers? Validating measures of effective teaching using random assignment*. Seattle, WA: Bill & Melinda Gates Foundation.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University.

Lampert, M. (2017). Ambitious teaching: A deep dive. In R. Heller, R. E. Wolfe, & A. Steinberg (Eds.), *Rethinking readiness: Deeper learning for college, work, and life* (pp. 147–173). Cambridge, MA: Harvard Education Press.

Lewis, C., & Tsuchida, I. (1999). A lesson is like a swiftly flowing river: How research lessons improve Japanese education. *Improving Schools*, *2*(1), 48–56.

National Governors Association Center for Best Practices (NGA Center) & Council of Chief State School Officers (CCSSO). (2010). *Common Core State Standards for Mathematics*. Washington, DC: Authors.

Nesmith, S. J. (2008). *Mathematics and literature: Educators' perspectives on utilizing a reformative approach to bridge two cultures*. Retrieved from https://files.eric.ed.gov/fulltext/EJ1099543.pdf

O'Day, J. A., & Smith, M. S. (2016). Quality and equality in American education: Systemic problems, systemic solutions. In I. Kirsch & H. Braun (Eds.). *The dynamics of opportunity in America*. Princeton, NJ: Educational Testing Service.

Pape, S. J., & Tchoshanov, M. (2001). The role of representation(s) in developing mathematical understanding. *Theory into Practice*, *40*(2), 118–127.

Perry, R., Marple, S., & Reade, F. (2017). *Instructional materials: Who makes the choice? Findings from the annual survey on implementing the Common Core State Standards in Mathematics*. San Francisco, CA: WestEd.

Perry, R., Marple, S., & Reade, F. (2019). *Roadblocks and routes: Professional development in Math in Common districts*. San Francisco, CA: WestEd.

Perry, R. R., Seago, N. M., Burr, E., Broek, M., & Finkelstein, N. D. (2015). *Classroom observations: Documenting shifts in instruction for districtwide improvement*. San Francisco, CA: WestEd.

Reade, F., Perry, R., & Heredia, A. (2018). *Improving the dynamics of classroom instruction in response to the Common Core: Results from the annual Math in Common teacher and administrator surveys*. San Francisco, CA: WestEd.

Reardon, S. F., Doss, C., Gagné, J., Gleit, R., Johnson, A., & Sosina, V. (2018). *Getting down to facts II: A portrait of educational outcomes in California*. Palo Alto, CA: Stanford University.

San Francisco Unified School District. (n.d.). *More resources for Math Talks*. Retrieved from http://www.sfusdmath.org/math-talks-resources.html

Schoenfeld, A. (2014). Classroom observations in theory and practice. *ZDM, the International Journal of Mathematics Education*, *45*, 607–621.

Schoenfeld, A. H., Floden, R. E., & Algebra Teaching Study and Mathematics Assessment Project. (2014). *The TRU math scoring rubric*. Berkeley, CA, & East Lansing, MI: Graduate School of Education, University of California, Berkeley, & College of Education, Michigan State University.

Seago, N., & Carroll, C. (2018). *Incremental shifts in classroom practice: Supporting implementation of the Common Core State Standards–Mathematics*. San Francisco, CA: WestEd.

Seago, N., & Perry, R. (2017). *Comparing features of high- and low-rated mathematics lessons*. Retrieved from https://www.wested.org/news-events/comparing-features-of-high-and-low-rated-math-lessons/

Smith, M., & Stein, M. K. (2011). *5 practices for orchestrating productive mathematics discussions*. Reston, VA: National Council for Teachers of Mathematics.

Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York, NY: The Free Press.

Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A., ... Ogbuehi, P. (2012). *Improving mathematical problem solving in grades 4 through 8: A practice guide* (NCEE 2012–4055). Retrieved from https://eric.ed.gov/?id=ED532215

Yoshida, M., & Jackson, W. (2010). *Improving blackboard organization to enhance students' mathematical thinking and understanding*. Retrieved from http://www.globaledresources.com/resources/assets/042210-Bansho.pdf

Appendix A. Observation Protocol: Details on the Eight Dimensions of Classroom Instruction

Table A1 provides additional details and source information on the eight dimensions on which WestEd observers gathered data. The first column includes the dimension name, the source, and (for MQI) the broad code area of the MQI instrument from which the dimension is derived. The second column includes information from the source instrument about the purpose of the dimension. The third column provides information on the features of an observed lesson that would be examined for evidence of the highest possible rating.

Table A1. Details and Source Information on the Eight Dimensions of Classroom Instruction

Dimension	Description or Guiding Questions	Characteristics of "High" or "Expert" Ratings
Linking representations [MQI; Richness of the Mathematics]	Teachers' and students' explicit, public (in small or whole groups) linking and connections between different representations of a mathematical idea or procedure. To count, these links must occur across different representational "families" (e.g., a linear graph and a table both capturing a linear relationship).	Links and connections related to student contributions are present, with extended, careful work characterized by one of the following features: • Explicit detail about how two or more representations are related (e.g., pointing to specific areas of correspondence) • Detail and elaboration about how two mathematical representations are related to each other
Multiple solutions/ procedures (MQI; Richness of the Mathematics)	Multiple procedures or solution methods occur or are discussed in the segment: • Multiple solution methods for a single problem (including shortcuts) • Multiple procedures for a given problem type	Multiple procedures or solution methods occur or are discussed in the segment, and include at least one of the following special features at some length: • Explicit extended comparison of multiple procedures or solution methods, for efficiency, appropriateness, ease of use, or other advantages and disadvantages • Explicit discussion of features of a problem that cues the selection of a particular procedure • Explicit links/connections between multiple procedures or solution methods (e.g., how one is like or unlike another)

Dimension	Description or Guiding Questions	Characteristics of "High" or "Expert" Ratings
Mathematical sense-making (MQI; Richness of the Mathematics)	The teacher publicly attends to one or more of the following: • The meaning of numbers • Understanding relationships between numbers • The relationships between contexts and the numbers or operations that represent them • Connections between mathematical ideas or between ideas and representations • Giving meaning to mathematical ideas • Whether the modeling of and answers to problems make sense	Teacher focuses on meaning in a sustained way during the segment. This need not be the entire lesson, but must be substantial.
Student explanations (MQI; Common Core–Aligned Student Practices)	Students provide a mathematical explanation for an idea, procedure, or solution. For example: • Students explain why a procedure works • Students explain the procedure they used to solve a particular problem by attending to the meaning of the steps involved in this procedure, rather than simply listing those steps • Students explain what an answer means • Students explain why a solution method is suitable or is better than another method • Students explain an answer based on an estimate or other number-sense reasoning	Student explanations characterize much of the lesson.

Dimension	Description or Guiding Questions	Characteristics of "High" or "Expert" Ratings
Student questioning and reasoning (MQI; Common Core–Aligned Student Practices)	Students engage in mathematical thinking that has features of important mathematical practices. Examples include, but are not limited to: • Students provide counterclaims in response to a proposed mathematical statement or idea • Students ask mathematically motivated questions requesting explanations (e.g., "Why does this rule work?") • Students make conjectures about the mathematics discussed in the lesson • Students form conclusions based on patterns that they identify or on other forms of evidence • Students engage in reasoning about a hypothetical or general case • Students use ideas from a different mathematical topic to reason about the content of the lesson • Students make a connection between the topic of the lesson and another mathematical area • Students comment on the *mathematics* of one another's contributions	Student mathematical questioning or reasoning characterizes much of the lesson.
The mathematics (TRU)	• How accurate, coherent, and well justified is the mathematical content (including mathematical language)? • Is there a clear mathematical goal for the lesson? • How did mathematical ideas develop within the lesson for students?	Classroom activities support meaningful connections among procedures, concepts, and contexts (where appropriate) and provide opportunities for building a coherent view of mathematics.
Access to mathematics (TRU)	• To what extent does the teacher support access to the content of the lesson for all students? • Who did and didn't participate in the mathematical work of the class, and how?	The teacher actively supports and, to some degree, achieves broad and meaningful *mathematical* participation OR what appear to be established participation structures result in such engagement.
Agency, authority, and identity (TRU)	• To what extent are students the source of ideas and discussion of them? • How are student contributions framed? • What opportunities did students have to explain their own and respond to one another's mathematical ideas? How does the teacher respond to student ideas?	Students explain their ideas and reasoning, and the teacher may ascribe ownership for students' ideas in exposition, AND/OR students respond to and build on one another's ideas.

Sources: Hill (2014) and Schoenfeld, Floden, & Algebra Teaching Study and Mathematics Assessment Project (2014).

Appendix B. Lesson Observation Sample

This table illustrates the number of lessons observed during each observation period in our study, by MiC district number and by grade level.

Table B1. Number of Lessons Observed, 2015–2018

District	Fall 2015	Spring 2016	Fall 2016	Spring 2017	Fall 2017	Spring 2018	Total
1	**4**	**6**	**5**	**6**	**0**	**0**	**21**
Elementary	2	3	1	3	0	0	9
Middle School	2	3	4	3	0	0	12
2	**6**	**6**	**5**	**3**	**0**	**6**	**26**
Elementary	3	3	3	2	0	4	15
Middle School	3	3	2	1	0	2	11
3	**7**	**7**	**6**	**5**	**5**	**4**	**34**
Elementary	6	6	4	4	3	2	25
Middle School	1	1	2	1	2	2	9
4	**0**	**5**	**0**	**6**	**0**	**6**	**17**
Elementary	0	1	0	4	0	4	9
Middle School	0	4	0	2	0	2	8
5	**4**	**6**	**4**	**3**	**3**	**0**	**20**
Elementary	0	0	2	2	2	0	6
Middle School	4	6	2	1	1	0	14
6	**6**	**6**	**5**	**5**	**5**	**4**	**31**
Elementary	3	3	3	3	2	2	16
Middle School	3	3	2	2	3	2	15
7	**0**	**0**	**4**	**6**	**0**	**0**	**10**
Elementary	0	0	2	3	0	0	5
Middle School	0	0	2	3	0	0	5

District	Fall 2015	Spring 2016	Fall 2016	Spring 2017	Fall 2017	Spring 2018	Total
9	**0**	**5**	**0**	**5**	**3**	**4**	**17**
Elementary	0	2	0	2	1	2	7
Middle School	0	3	0	3	2	2	10
10	**0**	**6**	**6**	**5**	**4**	**4**	**25**
Elementary	0	4	6	5	4	4	23
Middle School	0	2	0	0	0	0	2
Total	**27**	**47**	**35**	**44**	**20**	**28**	**201**

Appendix C. Statistical Analyses

The following tables illustrate analyses of the data from the 201 lessons that we observed, describing overall correlations among dimensions and underlying components that resulted in our groupings of connected dimensions in the report.

How are the eight dimensions of classroom instruction correlated with one another?

To address the question of how the eight dimensions of classroom instruction are correlated with one another, we computed the Pearson correlation coefficient. All 28 pairs of codes had Pearson correlation coefficients that were statistically significant, given the sample size of 201. We divided the resulting values into three groups, consisting of the pairs of codes that were correlated at relatively low ($\rho <= 0.38$, $n = 9$), medium ($0.38 < \rho < 0.53$, $n = 10$), and high ($\rho >= 0.53$, $n = 9$) levels, respectively. These three levels are displayed in Table C1 with color codes that indicate the strength of the correlation. (Table labels follow the same order as in Appendix A.)

The three TRU codes (shown as "Math," "Access," and "Agency" in Table C1) are correlated with one another at a high ($\rho \geq 0.54$) level. Several factors may contribute to this pattern:

- The TRU codes come from the same framework.
- The TRU codes have fewer rating levels (only three) than MQI, possibly

Table C1. How the Eight Dimensions of Classroom Instruction Are Correlated with One Another

	Mult	Sense	StExpl	StQandR	Math	Access	Agency
Linking	0.29	0.36	0.23	0.19	0.29	0.29	0.28
Mult		0.37	0.47	0.30	0.39	0.39	0.45
Sense			0.48	0.52	0.57	0.52	0.48
StExpl				0.56	0.54	0.55	0.62
StQandR					0.53	0.47	0.52
Math						0.60	0.55
Access							0.54

High   Medium   Low

Table C2. Underlying Components Incorporating Multiple Dimensions

	Component 1	Component 2	Component 3	Component 4	Component 5
Linking	0.1	0.97	0.11	0.13	0.12
Mult	0.24	0.12	0.93	0.13	0.16
Sense	0.13	0.22	0.21	0.78	0.38
StExpl	0.75	0.06	0.27	0.17	0.32
StQandR	0.67	0.01	-0.01	0.64	0.08
Math	0.3	0.09	0.15	0.42	0.67
Access	0.31	0.12	0.13	0.15	0.84
Agency	0.74	0.17	0.22	0.1	0.37

making it difficult to distinguish among lessons that actually differ substantively in quality.

- The "Apprentice" and "Expert" levels of ratings have a relatively low "floor."

Are there underlying components of multiple dimensions?

Principal components analysis is a data-reduction approach that can identify underlying components of individual codes or survey questions. There is a distinction drawn between constructs (which are created *a priori*) and components that are identified by the statistical computations. For MQI, one such construct is "richness of mathematics," which subsumes both the linking and multiple-solutions codes as well as codes for sense-making.

A principal components analysis with varimax rotation of the data set yielded the loadings with five underlying components, as shown in Table C2.

One interpretation of these results is that *linking representations* and *multiple solutions/procedures* remain codes that measure something distinct from the other codes. The MQI student-oriented codes (*student explanations* and *student questioning and reasoning*) are associated with the TRU *agency, authority, and identity* code. *Student questioning and reasoning* is also associated with *mathematical sense-making. The mathematics* and *access to mathematics* TRU codes are closely related to each other.

WestEd.org