# NCSM Journal

of Mathematics Education Leadership

SUMMER 2014 VOL. 15, NO. 2

National Council of Supervisors of Mathematics www.mathedleadership.org

# Table of Contents

# Analyzing Students' Work to Reflect on Instruction: The Instructional Quality Assessment as a Tool for Instructional Leaders

Melissa D. Boston, *Duquesne University*
Michael D. Steele, *University of Wisconsin-Milwaukee*

## Abstract

*In this article, we discuss how instructional leaders can use collections of students' work and the Instructional Quality Assessment (IQA) Mathematics rubrics to initiate conversations with groups of mathematics teachers and to monitor the success of professional development initiatives and curricular implementation efforts. In our work, collections of students' work are used to reflect on instructional practice, by considering the nature of instruction that supported students to produce the mathematical work and thinking. We ground the discussion in specific examples from two studies in which collections of students' work and the IQA rubrics were used to diagnose the effectiveness of professional development and curriculum implementation efforts, engage teachers in reflecting on practice, and inform next steps in the instructional change process.*

In the current era of the Common Core State Standards in Mathematics (CCSSM; Common Core State Standards Initiative [CCSSI], 2010) and increased accountability demands on teachers to support strong learning outcomes for all students, teachers and administrators are focused more closely than ever on the nature of teachers' classroom practice in mathematics (Cobb & Smith, 2008; Spillane, Halverson, & Diamond, 2004). Supporting students in meeting the Standards for Mathematical Practice (CCSSI, 2010) identified in the CCSSM will require a sharp departure from traditional procedurally driven mathematics curricula and teaching practices, and successful implementation of CCSSM will require "significant changes in the practice of most US mathematics teachers" (Cobb & Jackson, 2011, p. 185). To address these new demands, school districts across the country will need to engage teachers in professional learning experiences and adopt or revise mathematics curricula to promote the ambitious vision of mathematics teaching and learning advocated by the Standards for Mathematical Practice (Cobb & Jackson, 2011).

Instructional leaders play a critical role in the success or failure of teachers' efforts to grow and develop their classroom practice (Boyd et al., 2011; Tickle, Chang, & Kim, 2011). Studies of successful systemic change in secondary mathematics, for example, have identified strong instructional leadership as an integral component of changing classroom practice (e.g., Stein & Nelson, 2003; Stein, Silver, & Smith, 1998). Supporting meaningful change means that instructional leaders must engage with the substance of a reform initiative rather than simply the broad-stroke forms (November, Alexander, & van Wyk, 2010; Spillane, 2000; Stevens, 2004). While short walkthroughs and teacher observations are important tools that an instructional leader might use to support teacher professional development (Fink & Resnick, 2001), strong instructional leadership also includes engaging in conversations with teachers (individually and in professional learning communities) about instruction outside the context of an observation (Rossi, 2007).

Given the constraints on leaders' time, however, frequent full-class observations and debriefing sessions with teachers around classroom practice are often not feasible. Similarly, having groups of teachers observe each other may not be logistically possible to implement on a regular basis. As such, instructional leaders need tools that support meaningful discussions about teaching and learning with their mathematics department outside of regular class time. Conversations in professional leaning communities about the nature of mathematical tasks (e.g., Arbaugh & Brown, 2005) and the analysis of students' work (e.g., Kazemi & Franke, 2004) have been shown to be effective in supporting reflection on practice and teacher change. Interventions with principals have demonstrated that instructional leaders with diverse backgrounds can engage meaningfully in conversations about mathematics tasks, episodes of teaching, and students' work (Boston, Gibbons, & Henrick, 2011; Steele, Johnson, Otten, Herbel-Eisenmann, & Carver, under review). Each of these studies made use of research-based tools to structure conversations among teachers and administrators.

In this article, we present one such research-based tool – the Instructional Quality Assessment (IQA) – that can be used with student-work artifacts to analyze and interrogate the nature of classroom practice, as a proxy for classroom observations. We describe two projects in which students' work was collected as a measure of and reflection on instruction. In these settings, the IQA rubrics were used to analyze the effectiveness of a professional development initiative and the implementation of a standards-based, algebra curriculum. We suggest ways that instructional leaders could use the rubrics internally to serve diagnostic purposes and, most importantly, as a learning tool for fostering rich conversations with teams of mathematics teachers about mathematics instructional practice. While students' work has been used successfully to engage teachers in assessing students' thinking and understanding of mathematics, we propose that analyzing sets of students' work can also be used to initiate conversations about the nature of instruction that supported students to produce the mathematical work and thinking. In this way, students' work provides a reflection on instruction that can promote teachers' self-reflection, self–discovery, and transformative growth (Steele & Boston, 2012).

## *The Instructional Quality Assessment Mathematics Rubrics*

The Instructional Quality Assessment (IQA) Mathematics Toolkit was developed to provide a direct assessment of instructional quality based on live classroom observations or collections of students' work. Though initially created as a research instrument, the IQA can also serve as a tool to support rich conversations about instructional practices in mathematics. The IQA rubrics for classroom observations and students' work assess the rigor of instructional tasks, task implementation (i.e., how the demands of a task are enacted by teachers and students during instruction), classroom discourse (observation rubrics only), and teachers' expectations (students' work rubrics only). Research has consistently identified these four aspects of classroom instruction as impacting student achievement (Cobb, Boufi, McClain, & Whitenack, 1997; Hufferd-Ackles, Fuson, & Sherin, 2004; Staples, 2007; Stein & Lane, 1996). Figure 1 provides the *Teacher's Expectations* rubric and samples of teacher's expectations at each level. Excerpts from the rubrics for *Potential of the Task* and *Task Implementation* are provided in Figure 2 along with sample tasks and students' work indicative of the score levels on each rubric.[1]

The IQA rubrics are grounded in two bodies of research. First, the Mathematical Tasks Framework (Stein, Smith, Henningsen, and Silver, 2009) informed the IQA's assessment of instructional tasks separately from task implementation, and score levels within each rubric reflect the Levels of Cognitive Demand: *doing mathematics* and *procedures with connections* (i.e., high-level cognitive demands) and *procedures without connections* and *memorization* (i.e., low-level cognitive demands). Second, the collection and analysis of students' work as a valid reflection of instructional practice utilizes the research of Matsumura, Garnier, Pascal, and Valdes (2002). Design and generalizability studies determined that four sets of students' work, containing at least 4-6 samples per set and scored by two trained raters, provided a stable indication of a teachers' classroom practice highly correlated with observed instruction (Matsumura, Garnier, Slater, & Boston, 2008). As such, the analysis of samples of students'

[1] See Boston (2012) for a comprehensive description of the protocol for using the IQA rubrics and collecting and analyzing student work in research.

FIGURE 1: *IQA Mathematics Assignments rubric for* Teachers' Expectations *(Boston, 2012) and samples of* Teacher's Expectations

| | ***Teacher's Expectations* rubric** | **Samples of *Teacher's Expectations*** |
|---|---|---|
| 4 | The majority of the teacher's expectations are for students to engage with the high-level demands of the task, such as using complex thinking and/or exploring and understanding mathematical concepts, procedures, and/or relationships. | Sample 1 (for the Level 4 task in Figure 2): "I wanted to see students really thinking creatively about the problem, using what they know about benchmark fractions and percents, and using the diagram in their explanations. I wanted clear explanations that make sense to the reader." |
| 3 | At least some of the teacher's expectations are for students to engage in complex thinking or in understanding important mathematics. However, the teacher's expectations do not warrant a "4" because:<br>• the expectations are appropriate for a task that lacks the complexity to be a "4";<br>• the expectations do not reflect the potential of the task to elicit complex thinking (e.g., identifying patterns but not forming generalizations; using multiple strategies or representations without developing connections between them; providing shallow evidence or explanations to support conclusions).<br>• the teacher expects complex thinking, but the expectations do not reflect the mathematical potential of the task. | Sample 1 (for the Level 3 task in Figure 2): "To write a story problem that could be answered by solving the equation."<br><br>Sample 2: "I wanted students to be able to estimate perimeter and area and explain why they chose that particular estimate." [Teachers' expectations did not capture the main mathematical ideas of the task: developing students' understanding of perimeter and area by comparing the perimeter and area of irregular shapes.] |
| 2 | The teacher's expectations focus on skills that are germane to student learning, but these are not complex thinking skills (e.g., expecting use of a specific problem solving strategy, expecting short answers based on memorized facts, rules or formulas; expecting accuracy or correct application of procedures rather than on understanding mathematical concepts). | Sample 1 (for the Level 2 task in Figure 2): "My students always understand that quality work involves neatness, accuracy, and checks for accuracy. I continue to stress completeness, neatness, and accuracy… all problems attempted with minimal (1-2) mistakes."<br><br>Sample 2: "High performers were students who had math facts memorized and breezed through the assignment." |
| 1 | The teacher's expectations do not focus on substantive mathematical content (e.g., activities or classroom procedures such as following directions, effort, producing neat work, or following rules for cooperative learning). | Sample 1 (for the Level 1 task in Figure 2): "This work was checked for effort rather than performance. Students must label their papers (name, date, class) and use pencil (NOT pen)." |

work that conform to these requirements serves as a valid proxy for live classroom observation.

## *Using Student Work to Assess a Professional Development Initiative: The Summer Workshop in Mathematics Project*

During Summer 2010, the Summer Workshop in Mathematics (SWIM) project[2] engaged 39 elementary and middle school teachers (grades 3-8) in two, one-week professional development workshops focused on the teaching of fractions and algebraic thinking. The goal of the project was to develop teachers' understanding of fractions and/or algebraic ideas and to support teachers in analyzing the cognitive demands of mathematical tasks in any mathematical content area. As the primary professional learning activity in each workshop, teachers engaged in solving cognitively challenging tasks with the potential to engage them as learners in the Standards for Mathematical Practice and to support the development of their conceptual understanding of fractions and algebraic ideas. Each workshop also provided opportunities for teachers: to

[2] The first author served as Principal Investigator on the SWIM Project, funded by a grant from the Heinz Endowments.

FIGURE 2: *Excerpts of the* Potential of the Task and Task Implementation *rubrics from the IQA Mathematics Assignments Manual (Boston, 2012) and corresponding samples of students' work.*

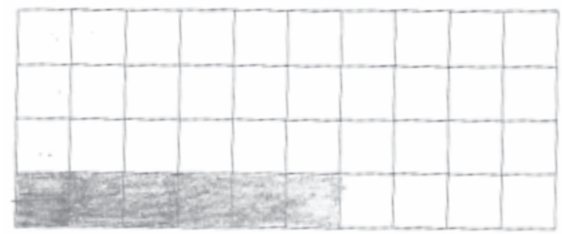
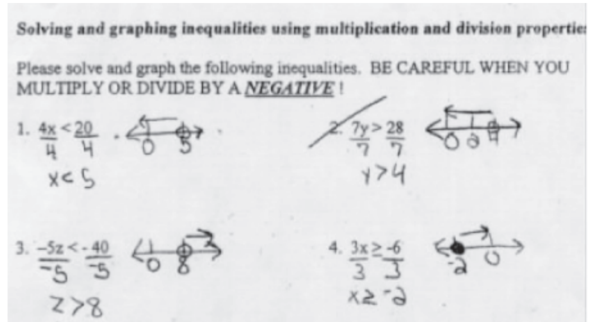
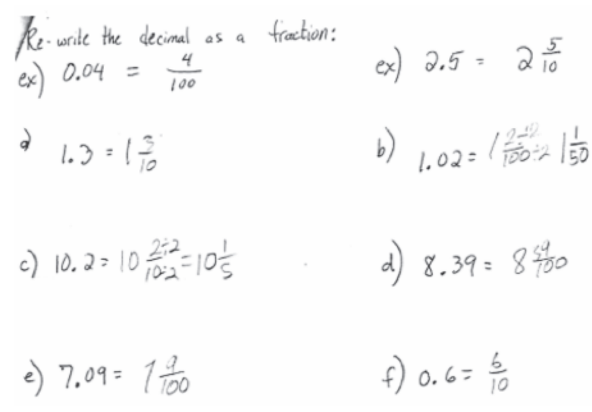
| | ***Potential of the Task* rubric** | ***Task Implementation* rubric** | **Sample of Student's Work** |
|---|---|---|---|
| 4 | The task has the potential to engage students in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:<br>• Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); or<br>• Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts.<br><br>The task *must explicitly prompt* for evidence of students' reasoning and understanding. For example, the task MAY require students to:<br>• solve a genuine, challenging problem for which students' reasoning is evident in their work on the task;<br>• develop an explanation for why formulas or procedures work;<br>• identify patterns; form and justify generalizations based on these patterns;… | Student work indicates the use of complex and non-algorithmic thinking, problem solving, or exploring and understanding the nature of mathematical concepts, procedures, and/or relationships (i.e., there is evidence of at least one of the descriptors of a "4" in the *Potential of the Task* rubric.) | Shade 6 of the small squares in the rectangle below:<br><br>Using the diagram, explain how to determine:<br>a) the percent of area that is shaded<br>First. The whole thing equals 100%. Then I knew that there are 4 rows so each row equals 25%. Then I divided 10 into 25 because there are 10 blocks in each row. My answer was 2.5 so if each block equals 2.5 I counted how many squares I shaded in and got 6. Then I multipled 2.5 x 6 and got 15 which is my answer |
| 3 | The task has the potential to engage students in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the task does not warrant a "4" because:<br>• the task does not explicitly prompt for evidence of students' reasoning and understanding.<br>• students may need to identify patterns but are not pressed to form or justify generalizations;<br>• students may be asked to use multiple strategies or representations but the task does not explicitly prompt students to develop connections between them;… | Student work indicates that students engaged in problem-solving or in creating meaning for mathematical procedures and concepts BUT student work lacks explicit evidence of complex thinking required for "4" (i.e., the *Potential of the Task* was rated as a 3 or 4…and there is a lack of evidence of the appropriate descriptors for a 4, but there is evidence of at least one descriptor of a 3). | Write a story problem for each of the following equations:<br>B) $144 = 24 + x$<br>Together Noah and Cory ate 144 pixi stix. Noah ate 24 pixi stix How many did Cory eat. |

FIGURE 2: *Excerpts of the* Potential of the Task *and* Task Implementation *rubrics from the IQA Mathematics Assignments Manual (Boston, 2012) and corresponding samples of students' work.*

| | ***Potential of the Task* rubric** | ***Task Implementation* rubric** | **Sample of Student's Work**<br>The potential of the task is limited to engaging stu- |
|---|---|---|---|
| 2 | dents in using a procedure that is either specifically called for or its use is evident based on prior instruction, experience, orplacement of the task. There is little ambiguity about what needs to be done and how to do it. The task does not require students to make connections to the concepts or meaning underlying the procedure being used… (e.g., practicing a computational algorithm). | Students engage with the task at a procedural level. Students apply a demonstrated or prescribed procedure. Students show or state the steps of their procedure, but do not explain or support their ideas. .. |  |
| 1 | The potential of the task is limited to engaging students in memorizing or reproducing facts, rules, formulae, or definitions… | Students engage with the task at a memorization level… (e.g., students provide answers only), OR even though a procedure is required or implied by the task, only answers are provided in students' work; there is no evidence of the procedure used by the students. |  |

compare tasks with different levels of cognitive demand (e.g., tasks that engaged students in reproducing procedures or memorized knowledge versus tasks that promoted reasoning, problem-solving and sense making); to analyze the cognitive demands of the tasks they engaged in solving; and to reflect on their experiences as learners and how the facilitator supported their learning. Beyond these reflections, however, teachers did not discuss the implementation of cognitively challenging tasks or how to enact the practices of the facilitator.

The questions guiding the study were: *Following the workshop, could teachers implement a high-level task in ways that maintained the cognitive demands, as evident in the sets of students' work? What were teachers' successes and challenges in maintaining high-level cognitive demands, as evident in the sets of students' work and expressed by teachers during the follow-up sessions?*

## Participants and Data

Thirteen teachers from Project SWIM elected to attend follow-up meetings during Fall 2010, to incorporate ideas from the workshops into their classrooms.[3] The teachers were from two urban school districts and two suburban school districts in a mid-sized Northeastern city. Teachers taught in 11 different elementary and middle schools, and all teachers had responsibility for teaching mathematics the majority of the school day. Demographic data for the teachers is provided in Table 1.

In the follow-up sessions, teachers used samples of students' work to describe their experiences in implementing high-level tasks, including successes and challenges. As samples of students' work were shared with the group, teachers could comment on what they noticed and wondered about students' mathematical understandings evident in the samples of work and the nature of the

[3] Other teachers declined participation in the follow-up sessions due to personal commitments, health issues, or teaching assignments in the new school year that did not include mathematics.

*Table 1: Demographic Data for Teachers in the Project SWIM Follow-Up Sessions*

| | | |
|---|---|---|
| School Setting: | Urban<br>11 | Suburban<br>2 |
| Age Level of Classroom (at time of project): | Elementary<br>(K-5)<br>7 | Middle School<br>(6-8)<br>6 |
| Teaching Certification: | Elementary<br>(K-6)<br>7 | Mathematics<br>(7-12)<br>6 |
| Gender: | Female<br>10 | Male<br>3 |

lesson in which the work was produced. Data from the follow-up sessions included teachers' written reflections (e.g., on instructional cases and on their own lessons and students' work), written artifacts produced during the follow-up sessions (i.e., chart paper listing successes and challenges in implementing high-level tasks), and the facilitator's notes from the discussions.

Ten of the 13 teachers agreed to provide their sets of students' work as data, resulting in 39 sets of student work for the analysis (four per teacher, with one teacher submitting only three). While the small sample size limits generalizations to the entire population of teachers in Project SWIM, the group of 10 teachers is important because 8 of them identified this type of mathematics instruction as atypical of their everyday practice. The sets of students' work captured their genuine, initial efforts to implement cognitively challenging instructional tasks, as might be the case in many districts embarking on instructional change in light of CCSSM.

## Analysis

Written artifacts, discussions, and teachers' reflections from the follow-up sessions were used to identify common successes and challenges in implementing cognitively challenging tasks, as reported by the teachers. Sets of students' work provided evidence of teachers' ability to maintain the cognitive demands of high-level tasks. Student-work sets were scored independently by two trained raters (the first author and a graduate research assistant not associated with the SWIM workshop), using the IQA Mathematics Assignments rubrics for *Potential of the Task* and *Task Implementation* (featured in Figure 2). The raters achieved 89% initial exact-point agreement, with all disagreements resolved through discussion. Consensus scores were used to produce descriptive statistics on the overall collection of students' work.

## Results

Within the group if 10 teachers, 8 (80%) teachers implemented at least 3 of 4 high-level tasks in ways that maintained students' opportunities for thinking and reasoning. In other words, the student-work samples provided evidence (as rated on the *Task Implementation* rubric) that students had actually engaged with the cognitively challenging aspects of the tasks (as rated on the *Potential of the Task* rubric). Overall, in 39 sets of students' work, 33 sets (85%) featured high-level tasks (i.e., a score of 3 or 4 on the *Potential of the Task* rubric) and 26 sets (67%) featured high-level implementations (i.e., a score of 3 or 4 on the *Task Implementation* rubric); hence 26 of 33 high-level tasks (79%) were maintained at a high-level during implementation. These data provide evidence that the majority of teachers were able to implement cognitively challenging instructional tasks.

Successes and challenges arose as teachers shared their experiences in implementing the tasks. Discussions among teachers regarding students' work samples served as a vehicle for identifying aspects of ambitious mathematics instruction that were present or absent from the student-work samples. Teachers often noticed successes as they examined other teachers' sets of students' work. For example, teachers commented that *students solved the task in more than one way even though the task directions did not specifically ask for multiple strategies, and students consistently used "because" in their written explanations.* These insights arose as teachers *noticed* aspects of other teachers' sets of students' work, and typically generated discussion as they *wondered* how students had been 'trained' to solve the task in more than one way or include a conceptual explanation even when not explicitly prompted by the task (i.e., how these norms had been developed in the classroom).

Challenges arose in teachers' reflections and noticings on their own and other teachers' sets of students' work. Common challenges included: resources for high-level tasks (noticing that the task was not high-level); evidence of students' lack of a conceptual understanding (noticing that students could not solve the cognitively challenging aspects of the task); and the quality of written explanations (identified by a teacher regarding his/her own students' work, and relating to the low-quality of verbal explanations during the lesson). Challenges regarding teachers' own student-work sets were sometimes noted by the teacher initially, and sometimes arose in comparison to other teachers' student-work. For example, the challenge of improving students' verbal and written explanations was identified by several teachers in the first follow-up meeting. Teachers noted students' difficulty in providing verbal explanations during class, and how this was evident in students' written explanations on the student-work samples. While reviewing a set of students' work from another teacher, one teacher reflected back to her own students' explanations: "Even though students were writing 'explanations,' the explanations only involved procedural steps." A discussion ensued regarding the difference between procedural and conceptual explanations, and how to develop students' ability to create conceptual explanations, especially in classrooms where cognitively challenging mathematical work and thinking were new experiences for teachers and students. Teachers collectively took this issue on as a group, brainstormed ideas, and returned to the second follow-up session eager to share new instructional practices (e.g., having a student provide a verbal explanation as another student writes what is being said, then both students revise the explanation to clearly communicate the mathematical thinking; prompting students to explain their thinking as if they were talking or writing to someone in a younger grade or a classmate who was absent from the lesson).

### Implications for Instructional Leaders

In addition to identifying successes and challenges in implementing tasks, the student-work collection also served diagnostic purposes, identifying successes of the professional development initiative and pathways for improvement for the teachers as a group (analogous to instructional leaders using student-work diagnostically within a department, school, or district). First, after participating in the professional learning experiences of solving cognitively challenging tasks, participants appeared to be successful in selecting high-level tasks (85% of tasks were high-level) and in supporting students' exploration of cognitively challenging instructional tasks (79% of tasks that began as high-level were maintained during instruction). These sets of students' work provided evidence that students solved tasks in a variety of ways, and used manipulatives, diagrams, and representations to support their thinking. Samples of student-work had unique strategies and ways of thinking, and did not look uniform (i.e., as though students had been directed on how to solve the tasks). If the teachers were within the same school or district, the instructional leader would want to capitalize on the fact that most teachers could identify and implement a high-level task successfully, and base future professional development initiatives on this foundation.

Second, high-level task demands that declined during implementation, even from a *Potential of the Task* score of 4 (i.e., the task explicitly required explanations of students' high-level work and thinking) to a *Task Implementation* score of 3 (i.e., implicit evidence of students' high-level thinking), could often be attributed to non-existent or low-quality written explanations. This indicates that, while teachers' experiences solving cognitively challenging tasks as learners enabled them to implement high-level tasks in ways that encouraged multiple strategies and representations, teachers did not appear to gain ways of developing students' mathematical explanations. As a next step, instructional leaders would want to provide opportunities for professional learning experiences specifically focused on supporting students to clearly explain their thinking, verbally and in writing.

Third, teachers with curricula lacking in high-level tasks often used open-ended assessment items or tasks directly from the workshop for their student-work collections. As teachers identify the need for curricular materials containing high-level instructional tasks, an instructional leader would want to provide teachers with increased access to curriculum and resources containing such tasks. However, research cautions that simply providing teachers with new or revised curricular materials does not guarantee that the materials will be implemented as intended (e.g., Remillard & Bryans 2004). In the next section, we discuss how instructional leaders can use collections of students' work to diagnose and support teachers' implementation of a cognitively challenging mathematics curriculum.

## *Using Student Work to Assess Curriculum Implementation: The Mathematical Practices Implementation Study*

Another approach to supporting instructional change involves implementing new or revised mathematics curricula. The success of such an implementation presents a number of challenges for administrators, teachers, and students. At the high school level in particular, curricula that feature an abundance of high cognitively demanding tasks and support discourse-based pedagogies require substantial systemic support, and even with that support such curricula often lose traction when key personnel leave the district (Senk & Thompson, 2003; St. John et al., 2005). The Mathematical Practices Implementation (MPI) study is analyzing the implementation of one such curriculum, the Education Development Center's "Center for Mathematics Education" (CME) Project.[4] The goals of the study are to measure the extent to which implementation of the CME Project materials reflected the high cognitive demands of the curriculum, and to identify key factors that support or inhibit the principled implementation of the curriculum. By principled implementation, we mean teaching that is faithful to the overarching principles and mathematical habits of mind upon which a curriculum is built (Cuoco, Goldenberg, & Mark, 1996), moving beyond simpler measures of textbook use to capture the ways in which the curricular tools are used in teaching. To understand the extent to which teaching represents principled implementation, the MPI study seeks to measure a number of aspects of teaching practice, including teachers' mathematical knowledge for teaching, their understanding of the mathematical habits of mind, the influence of teacher professional development on implementation, and the ways in which classroom norms and practices support student engagement and learning. This analysis considered specifically the relationships between the potential of the tasks teachers select for students to work, the implementation of those tasks as measured by the student work, and the expectations teachers have for the work students will produce on those tasks.

## Participants and Data

We identified two large metropolitan districts that were adopting CME Algebra I at the start of the study, and recruited 50 teachers at 12 school sites to participate in the study. These twelve school sites were housed in ten districts across five states, in or adjacent to urban centers serving a diverse student population. Teachers ranged in experience from 0 to more than 20 years of experience (see Table 2); 98% held a secondary mathematics certification, with the remaining 2% holding a certification in a secondary field other than mathematics. Teacher-participants at each site committed to submitting four days' worth of assignments completed during class time twice a year (fall and spring) across the first two years of implementation.

## Analysis

Project personnel scored the student-work samples using the IQA rubrics for *Potential of the Task, Task Implementation*, and *Teacher's Expectations* (Figures 1 and 2). Raters that demonstrated 85% agreement or better on test items rated the project data samples. The project also assessed the academic rigor of the curriculum materials, rating each section of the Algebra I text using the IQA *Potential of the Task* rubric. These ratings were used to compare the potential of the specific tasks teachers selected for students to the section's potential in general (i.e., were teachers selecting the high-level tasks available in each section of the curriculum?). The study is presently at the end of its first year of data collection, with the first two sets of students' work rated for participating teachers. At present, the first year data set contains 85 discrete student work sets, which were analyzed for this study.

## Results

Two important trends emerged from the student work ratings thus far that have implications for the support of a new curriculum implementation. The first trend relates to the potential of the tasks that teachers implemented with their students. Across the data set, the tasks teachers used with students almost universally reflected a lower *Potential of the Task* rating as compared to the text sections to which the assignments corresponded. Of the 85 student work samples rated, 67% were rated as a *Potential* of 2, indicating that students executed a clear mathematical procedure without providing implicit or explicit connections to meaning. This indicated that teachers in their first year implementing the new curriculum overwhelmingly selected

*Table 2: Years of Experience for MPI Study teachers, Year 1*

| | |
|---|---|
| 0 years (first year teaching) | 8% |
| 1 year | 10% |
| 2-5 years | 28% |
| 6-10 years | 28% |
| 11-15 years | 18% |
| 16-20 years | 2% |
| More than 20 years | 6% |

[4] The second author serves as a Co-Principal Investigator on the MPI Study, funded by the National Science Foundation.

procedural tasks, choosing not to implement higher cognitively demanding tasks that asked students to make sense of the underlying mathematical ideas. While this finding may be disappointing, it is not necessarily unexpected given prior research regarding teachers' selection of high and low cognitively demanding tasks (e.g., Stein & Lane, 1996). We also noticed that the bulk of high cognitively demanding tasks that teachers selected declined in rigor with respect to implementation and sought to identify reasons that might explain these declines.

This investigation led to our second finding, which was particularly illuminating with respect to principled implementation. For the 22 tasks that began at a high level (3 or 4 on the IQA scale for *Potential of the Task*) and declined in implementation (1 or 2 on the *Implementation* scale), we also looked at the scores for the teacher's expectations for the assignment. Scores of 1 and 2 on the *Teacher's Expectations* rubric represent expectations that are either non-mathematical in nature, such as neatness or clarity, or that are not complex thinking skills, such as short answers or accurate application of procedural steps. Although the rigor of these expectations is appropriate for tasks of a low potential, they are not a good fit for tasks of higher potential. We defined tasks with a *Potential* score of 3 or greater and an accompanying set of *Teacher's Expectations* scores of 2 or less as *Potential-Expectation* mismatches. If the *Potential* and *Teacher's Expectations* were both low (2 or less) or both high (3 or greater), we identified this as a *Potential-Expectation* match.

Across the first year data set, 86% of tasks that declined from high to low cognitive demand also featured a *Potential-Expectation* mismatch. Only in 3 of 22 cases did the implementation of a task represent complex thinking despite a lack of explicit expectations for complex thinking as measured by the rubric. This suggested an important relationship between the rigor of the teacher's expectations and students' engagement with the task: a low rigor of expectation engenders student work that systematically does not attend to the high cognitive demand aspects of the task.

There were also some promising signs of change in the first year of the study. Between the fall and spring data collections, the average score on the *Teachers' Expectations* rubric rose from 1.65 to 1.96. This suggested a trend in which the rigor of teachers' expectations increased over the course of the first year of implementation. This finding also indicated that explicit conversations about the ways in which the expectations can support students' engagement in high cognitively demanding tasks (perhaps through emphasis on the mathematical habits of mind) might further support teachers in moving towards a principled implementation of the CME Project Algebra I curriculum. Supporting teachers in being able to describe and set expectations for rich mathematical thinking is also likely to support students in being more successful in engaging in the Standards for Mathematical Practice, a key aspect of the new multi-state assessment systems.

## Implications for Instructional Leaders

Translating curricular resources into instruction that results in deep student learning can be a challenging task for teachers, particularly with curricula like the CME Project that support ambitious visions of teaching. This preliminary analysis of student work from teachers in their first year of implementing CME Algebra I suggest some specific ways in which instructional leaders might support such a curriculum implementation. The first area of support is the selection of tasks from a section of the text in which to engage students. To support teachers in selecting tasks that better represent the cognitive demand of a given text sections, instructional leaders and teachers might work on the mathematical tasks in the section together and discuss ways in which to support students in thinking through high cognitively demanding tasks. Particularly for districts that are moving from curricula with a heavier skills emphasis, teachers may be more disposed to select the familiar procedural tasks from a text section. Discussing the task selection process with teachers and understanding their decision-making process could help instructional leaders support a long-term systemic implementation of ambitious curricula.

Second, instructional leaders and teachers might find a benefit in co-designing expectations for students that support high cognitively demanding work. This work could be done either with respect to specific mathematics content, or in the form of general rubrics that teachers might apply across a broad range of student work. Working with teachers to set and communicate these expectations can help to send important messages to students that thinking and reasoning is a valued part of their mathematical work, rather than simply correct answers or properly executed procedures.

In the next section, we generalize across the two specific studies presented herein to discuss how instructional leaders might use collections of students' work and the IQA rubrics to support their work more broadly.

## *Using Student Work in Instructional Leadership: Initiating Conversations and Diagnosing Next Steps*

The specific studies discussed herein represent two situations of instructional change that will be common to many districts embarking on successful implementation of the CCSSM: professional development to support teachers' initial attempts in using cognitively challenging instructional tasks, and the implementation of new or revised curricula consisting primarily of cognitively challenging instructional tasks. In this article, we have presented how the analysis of sets of students' work can be used to diagnose the success of each of these efforts and provide instructional leaders with data to inform instructional improvements. As applicable beyond our specific studies, mathematics teachers participating in instructional change efforts (i.e., sustained professional development experiences or curriculum implementation) consistent with CCSSM can be asked to collect samples of students' work to tell the story of their successes and challenges in implementing cognitively challenging instructional tasks. Across a school or district, collections of students' work can then serve as artifacts for instructional leaders to initiate conversations about instructional practice amongst teachers and to identify pathways for instructional improvements.

### Initiating Conversations

In our work, we have utilized a variety of formats in leveraging students' work to initiate conversations with teachers. One method is to use an open case story format (Hughes, Smith, Boston, & Hogel 2008), where teachers share their experiences implementing high-level tasks, use the student-work samples as evidence of their successes and challenges, and allow other teachers to share their noticing and wonderings. Another method is to explicitly use the IQA rubrics to guide the conversation. Teachers can use the *Potential of the Task* rubric to identify the level of cognitive demand of the task and to identify specific aspects of the task that make it high-level. They are then asked to consider, "What is the evidence that students engaged with the high-level demands of the task?" and to use this evidence to score the set of students' work (holistically) based on the IQA *Task Implementation* rubric. Similarly, teachers can be asked to compare the score for *Potential* of the tasks used for instruction with either a) the *Potential* of the tasks featured in the corresponding section of the curriculum, to determine whether they are capitalizing on the cognitively challenging tasks featured in the curriculum or b) the rigor of their *Expectations* for the task, to identify *Potential-Expectation* matches or mismatches. Since teachers are looking across a set of responses rather than at individual student's work and thinking, commonalities often arise that can be attributed to the nature of instruction rather than to the students' mathematical thinking or ability.

In this way, reflecting on students' work (from their students and from other teachers' students) serves a self-diagnostic purpose, where teachers identify aspects of instructional practice that support or inhibit students' opportunities to engage in high-level thinking and reasoning. Instructional leaders can use collections of students' work to initiate similar conversations with teachers, where insights for instructional improvement are identified by the teachers themselves.

### Identifying Pathways for Instructional Improvement

Analyzing collections of student work can also serve diagnostic purposes for instructional leaders, by considering, "What does the collection of students' work indicate about the quality of instruction and students' learning opportunities in my department, school, or district?" Using the IQA rubrics specifically, instructional leaders can use collections of students' work to address questions about:

- **Instructional tasks:** Are teachers using high-level tasks for instruction? Are teachers choosing the high-level instructional tasks featured in their curriculum?
- **Task implementation:** Are teachers implementing instructional tasks in ways that maintain the cognitive demands and mathematical purposes of the tasks? What are teachers' specific successes and challenges in implementing high-level tasks?
- **Classroom norms and practices:** What opportunities do students have to demonstrate and explain their mathematical work and thinking in writing? What representations are students provided opportunities to use? What counts as an explanation?
- **Teachers' expectations:** What is the level of rigor of teachers' expectations for students' mathematical work

and thinking? Are these expectations aligned with the tasks embedded in the curriculum?

In this way, analyzing collections of student work supports instructional leaders in identifying pathways for instructional improvement.

Of course, there are important questions regarding the quality of classroom instruction that collections of students' work cannot answer. For example, what types of questions are being asked? What is the quality of mathematical discourse? How do the teacher-student and student-student interactions support students' mathematical learning? These questions might best be addressed through focused classroom observations and could complement discussions around students' work.

## *Conclusion*

Many school districts will need to implement professional development initiatives and new or revised curricula to enable teachers and students to meet the expectations of the Standards for Mathematical Practice (Cobb & Jackson, 2011). As supporting instructional change in the mathematics classroom is particularly challenging, especially with secondary teachers, school leaders will need to take an active role to guarantee the success of these efforts (Stein & Nelson, 2003). The analysis and discussion of collections of students' work using the IQA rubrics can provide principals, curriculum supervisors, and department chairs with reliable, valid, and mathematically rigorous tools to engage teachers in discussions and collect diagnostic data to monitor change. The IQA students' work rubrics provide valuable tools for instructional leaders by: identifying aspects of instructional practice that matter in terms of students learning; identifying areas of instructional improvement at the scale of a school or district; aligning well with the Standards for Mathematical Practice; and being well-suited for assessing professional development and curriculum implementation efforts. Moreover, the collection of students' work represents a way of discussing instructional practice that does not hinge on the availability for observations, provides a permanent artifact of practice that can easily be shared for analysis and discussion within a professional learning community, and is easier to collect and less intrusive to teaching than group observations or videotaping. Students' work also backgrounds the teacher's actions, making it a safer space (than video or observation) to discuss the successes and challenges encountered in one's own classroom. Within group of teachers, collections of students' work can serve as evidence of practice that is not subjective or reliant on teachers' recapping of events in a lesson, which can help to avoid judgment or debates over what happened. Most critically, discussing students' work from teachers' own classrooms positions teachers as decision makers in the process of instructional change, providing opportunities for collective- and self-reflection on teaching practice. Since the work was generated by students within their own school, district, or region, teachers often come to realize the mathematical capabilities of their own students by analyzing students' work from other teachers' classrooms. Collectively, teachers identify common issues and challenges in enacting ambitious instruction and construct pathways for improving their practice in ways that better support students' learning. ✪

## *References*


Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: A catalyst for change? *Journal of Mathematics Teacher Education, 8*, 499-536.

Boston, M. D., Gibbons, L., & Henrick, E. (2011, April). *Using classroom observation instruments to improve principal's capacity as instructional leaders in mathematics.* Paper presentation at the annual meeting of the American Education Research Association, New Orleans, LA.

Boston, M. D. (2012). Assessing the quality of mathematics instruction. *Elementary School Journal, 113*, 76-104.

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The influence of school administrators on teacher retention decision. *American Educational Research Journal, 48*, 303-333.

Cobb, P., & Jackson, K. (2011). Towards an empirically grounded theory of action for improving the quality of mathematics teaching at scale. *Mathematics Teacher Education and Development, 13*, 6-33.

Cobb, P., & Smith, T. M. (2008). District development as a means of improving mathematics teaching and learning at scale. In K. Krainer & T. Wood (Eds.), *Participants in mathematics teacher education: Individuals, teams, communities, and networks* (Vol. 3, pp. 231-254). Rotterdam, The Netherlands: Sense Publishers.

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education, 28,* 258-277.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics.* Washington, DC: National Governors Association Center for Best Practices and Council of Chief State School Officers.

Cuoco, A., Goldenberg, E. P., & Mark, J. (1996). Habits of mind: An organizing principle for mathematics curriculum. *Journal of Mathematical Behavior, 15*, 375-402.

Fink, E., & Resnick, L. B. (2001). Developing principals as instructional leaders. *Phi Delta Kappan, 82,* 598-606.

Hufferd-Ackles, K., Fuson, K., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education, 35*, 81-116.

Hughes, E. A., Smith, M. S., Boston, M. D., & Hogel, M. (2008). Case stories: Supporting teacher reflection and collaboration on the implementation of cognitively challenging mathematical tasks. In F. Arbaugh & P. M. Taylor (Eds.), *Inquiry into mathematics teacher education* (pp. 71 - 84). San Diego, CA: Association of Mathematics Teacher Educators.

Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education, 7*, 203-235.

Matsumura, L. C., Garnier, H., Slater, S. C., & Boston, M. D. (2008). Toward measuring instructional interactions "at-scale." *Educational Assessment, 13*, 267-300.

Matsumura, L. C., Garnier, H., Pascal, J., & Valdés, R. (2002). Measuring instructional quality in accountability systems: Classroom assignments and student achievement. *Educational Assessment, 8*, 207-229.

November, I., Alexander, G., & van Wyk, M. M. (2010). Do principal-educators have the ability to transform schools? A South African perspective. *Teaching and Teacher Education, 26*, 786–795.

Remillard, J. T., & Bryans, M. B. (2004). Teachers' orientation toward mathematics curriculum materials: Implications for teacher learning. *Journal for Research in Mathematics Education, 35*, 352-388.

Rossi, G. A. (2007). The classroom walkthrough: The perceptions of elementary school principals on its impact on student achievement (Unpublished doctoral dissertation). University of Pittsburgh, Pennsylvania.

Senk, S. L., & Thompson, D. R. (2003). *Standards-based mathematics curricula: What are they? What do students learn?* Mahwah, NJ: Lawrence Erlbaum.

Spillane, J. P. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction, 18*, 141–179.

Spillane, J. P., Halverson, R., & Diamond, J. B. (2004). Towards a theory of leadership practice: A distributed perspective. *Journal of Curriculum Studies, 36*, 3-34.

Staples, M. (2007). Supporting whole-class collaborative inquiry in a secondary mathematics classroom. *Cognition and Instruction, 25*, 161-217.

Steele, M. S., & Boston, M. D. (2012, April). *Using student work as a reflection on instruction.* Paper presentation at the Research Pre-session of the National Council of Teachers of Mathematics Annual Meeting, Philadelphia, PA.

Steele, M. D., Johnson, K., Otten, S., Herbel-Eisenmann, B. A., & Carver, C. (under review). Learning about algebra teaching, learning, and leading: Engaging principals in content-based professional development. *Journal of Research on Leadership Education.*

Stein, M., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*, 50-80.

Stein, M. K., & Nelson, B. S. (2003). Leadership content knowledge. *Educational Evaluation and Policy Analysis, 25*, 423–448.

Stein, M. K., Silver, E. A., & Smith, M. S. (1998). Mathematics reform and teacher development: A community of practice perspective. In J. G. Greeno & S. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 17–52). Hillsdale, NJ: Erlbaum.

Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2009). *Implementing standards-based mathematics instruction: A casebook for professional development* (2nd ed.). New York, NY: Teachers College Press.

Stevens, R. J. (2004). Why do educational innovations come and go? What do we know? What can we do? *Teaching and Teacher Education, 20*, 389–396.

Tickle, B. R., Chang, M., & Kim, S. (2011). Administrative support and its mediating effect on US public school teachers. *Teaching and Teacher Education, 27*, 342–349.