Developing Visions of High-Quality Mathematics Instruction

Charles Munter

University of Pittsburgh

Author Note

This material is based upon work supported by the National Science Foundation under Grant No. ESI 0554535, Paul Cobb and Thomas Smith, co-PIs, and by the Institute of Education Sciences, U.S. Department of Education, through grant R305B040110 to Vanderbilt University. The author wishes to thank Anne Garrison Wilhelm for her assistance with developing and managing the coder agreement process; Anne Garrison Wilhelm, Lynsey Gibbons, Kara Jackson, and Christine Larson for their assistance in refining the rubrics and coding interviews; Dan Berebitsky, Glenn Colby, Angela Hacker, Adrian Larbi-Cherif, Brooks Rosenquist, Rebecca Schmidt, and Jonee Wilson for coding interviews; Paul Cobb for reading and commenting on previous drafts; Rogers Hall for early conceptual guidance; and Peg Smith, the editor, Cynthia Langrall, and the 11 anonymous reviewers for additional comments and suggestions.

Correspondence concerning this paper should be addressed to Chuck Munter, University of Pittsburgh, Department of Instruction and Learning, 230 S. Bouquet St., 5300 Wesley W. Posvar Hall, Pittsburgh, PA 15260. Email: cmunter@pitt.edu, Phone: (412)648-1079.
Abstract

This paper introduces an interview-based instrument that was created for the purposes of characterizing teachers’, principals’, mathematics coaches’, and district leaders’ visions of high-quality mathematics instruction (VHQMI) and tracking changes in those visions over time. The instrument models trajectories of perceptions of high-quality instruction along what have been identified in the literature as critical dimensions of mathematics classroom practice. Included are a description of the methods by which an analysis of interview data were integrated with previous findings from the research literature in order to develop leveled rubrics for assessing VHQMI, a report of the results of using the instrument to code more than 900 interviews, and a discussion of the possible applications and benefits of such a methodological approach.
Developing Visions of High-Quality Mathematics Instruction

During the last two decades the mathematics education community has made great strides in uncovering and documenting important nuances of the phenomena it studies. And, the community has progressed a great deal in developing a language for naming and documenting the complex practices of high-quality mathematics instruction. However, while most of the knowledge we have accumulated has painted an increasingly clear portrait of mathematics instruction that gives all students access to significant mathematical ideas, the paths that teachers’ learning might take as they develop such instructional practices have not been characterized with equal clarity. As Staples (2007) argued, based on her work on collaborative classrooms, research has focused on describing instructional practices “only once the practices have been established” (p. 164, italics in original). Furthermore, the language the research community has developed for describing these aspects of classroom activity has grown primarily from etic accounts, often ignoring both participants’ ways of conceiving of and describing instructional practices and the influences of the school and district settings in which they work (Cobb, McClain, Lamberg & Dean, 2003; Hammerness, 2001).

Recently, members of the Middle School Mathematics and the Institutional Setting of Teaching (MIST) project collaborated for four years with four urban school districts that had set ambitious goals for reforming mathematics instruction (Cobb & Smith, 2008). Leaders in each district had a vision for classroom mathematics instruction and were working to support teachers' development of instructional practices consistent with that vision, which is justified in terms of day-to-day classroom learning opportunities for children. Hence, the MIST project provided an opportunity to investigate teachers’ and others’ evolving conceptions of high-quality mathematics instruction. The research question that guided my work was: What trajectories do
mathematics teachers’ and others’ visions of high-quality mathematics instruction follow over time in settings in which leaders are promoting models of instruction aligned with mathematics education research?

In this paper, I describe an interview-based assessment instrument developed in the course of the MIST project for characterizing teachers’, principals’, mathematics coaches’, and district leaders’ visions of high-quality mathematics instruction and tracking changes in those visions over time. The instrument consists of a series of interview prompts and accompanying rubrics, which model trajectories of perceptions of high-quality instructional practice along what have been identified in the literature as critical dimensions of mathematics classroom teaching and learning.

**Visions of High-Quality Mathematics Instruction**

**Background and Rationale**

In its first funding cycle, the MIST research team worked for four years (2007–2011) with each of four urban school districts serving ethnically and economically diverse populations as co-designers of support structures and strategies for meeting ambitious goals for reforming mathematics instruction at a district level. At the outset of the study, all four participating districts had recently formulated and begun implementing comprehensive initiatives promoting mathematics classrooms in which all middle grades students regularly collaborate to make sense of and solve challenging mathematical tasks and engage in rigorous discussions about their thinking. These efforts included the provision of professional development and curricula intended to support teachers in developing instructional practices that would afford such opportunities. The MIST goal was to investigate, test, and refine a set of conjectures about how multiple aspects of the instructional settings in which these reforms were being undertaken—
including organizational structures, social relationships, and material resources—enhance the impact of professional development on middle school mathematics teachers’ instructional practices.

Of course, among the myriad initiatives, structures and relationships that teachers experience are potential sources of incoherence or incompatibility, including competing agendas or merely conflicting interpretations. As Spillane, Reiser, and Gomez (2006) found, individuals working in various district units can interpret district reform initiatives for mathematics instruction differently. To the extent that district leaders’ interpretations and practices are in conflict, they can pose problems in terms of coherence in district-wide instructional improvement efforts and the degree to which teachers are supported in improving their instructional practices (Cobb & Smith, 2008). As the results of a number of studies have revealed, professional development, collaboration between teachers, and collegiality between teachers and school leaders are rarely effective unless they are tied to a “shared vision” of high-quality instruction that gives them meaning and purpose (Peterson, McCarthey, & Elmore, 1996; Newman & Associates, 1996; Rosenholtz, 1985, 1989; Rowan, 1990).

One of the MIST project’s initial conjectures, then, was that improvement in teachers’ instructional practices would be greater in schools where teachers and instructional leaders share a vision for high-quality mathematics instruction. Our assumption was not that a shared vision is something that individuals decide to have, but something that is built over time as colleagues’ professional discourses change (and converge). Thus, in order to assess the extent to which participants’ instructional visions are compatible with those of colleagues and the vision promoted by the district, we needed a means of tracking shifts in leaders’ and teachers’ articulations of “what it is important for students to know and be able to do mathematically
and… how students’ development of these forms of mathematical knowledgability can be effectively supported” (Cobb & Smith, 2008, p. 7).

The work described in this report is a first step toward meeting that need. It does not yet provide a means of assessing the sharedness of instructional visions; rather, it is motivated by the need to first understand individual mathematics teachers’ and others’ perceptions of high-quality instructional practice and what they might look like in transition. Its aims are similar to those of Simon, Tzur, Heinz, Kinzel, & Smith (2000), who, based on their study of 19 pre- and in-service teachers whose practices were transitioning in response to reform and their work in university teacher education courses, identified distinct perspectives of mathematics and mathematics learning that teachers might hold as their understandings, beliefs, or practices evolve. As those authors noted, efforts to identify and articulate such milestones respond to Goldsmith and Schifter’s (1997) still-applicable critique that “[t]he motivation for helping teachers develop new forms of practice is high, but the means by which teachers actually do so are currently not well understood” (p. 20).

Although developmental characterizations of teachers’ and others’ (visions of) practice are unfortunately rare, a few precedent cases exist. For example, researchers of Cognitively Guided Instruction proposed a 5-level trajectory of cognitively guided instructional practices pertaining to eliciting and making use of children’s thinking, which described milestones in teachers’ evolving practices (Fennema et al., 1996; Franke, Carpenter, Levi, & Fennema, 2001). As a precursor to those efforts, Peterson, Fennema, Carpenter, and Loef (1989) proposed a framework for characterizing teachers’ changing pedagogical content beliefs as more or less cognitively based, for which interview responses were rated on a 1-5 scale with respect to each of four constructs. Similarly, Simon and Schifter (1991) developed the Assessment of
Constructivism in Mathematics Instruction to assess teachers’ development of constructivist views of learning.

To better understand teacher learning and how to support professional growth, I contend that such documentation of the development of (visions of) practice is needed across all relevant dimensions of classroom instruction. The mathematics education research community may be articulating increasingly clear end goals for teachers’ knowledge and practice, but without sufficient “roadmaps,” how can those working to support teacher learning know whether progress is being made? Thus, with the rubrics described in this article, my goal was to fill in what I perceived as missing pieces in the literature by modeling more complete trajectories from less to more sophisticated instructional visions.

**Defining “Visions” (of High-Quality Mathematics Instruction)**

The instructional visions that I have worked to characterize among teachers and others are similar to Hammerness’s (2001) notion of *teachers’ vision*, which she defined as “a set of images of ideal classroom practice for which teachers strive” (p. 143). She described the role that teachers’ personal visions can play in their own learning and argued that such visions should be foundational in school improvement efforts. Through interview-based analyses of 16 cases, Hammerness demonstrated how different categories of visions—“constellations” of variance with respect to *focus* (the primary images or ideas on which a vision is concentrated), *range* (the extent to which a vision extends beyond one aspect of practice), and *distance* (the discrepancy between one’s vision and current practice)—related to the contexts in which teachers were working. “Acknowledging and surfacing teachers’ personal visions,” she argued, “might assist reformers to move beyond thinking about teachers simply as ‘advocates’ and ‘skeptics’ to a deeper understanding of teachers’ response to reforms” (p. 160). Like Hammerness, I employ the
term *vision* rather than *beliefs*. More than avoiding a construct for which no clear definition has emerged (Philipp, 2007), I am attempting to make an important semantic distinction that responds both to the data collected in interviews with study participants, and to ways of seeing the world that encompass horizons not yet reached. Whereas “beliefs” suggest a relatively static set of decontextualized ontological commitments, “vision” is intended to communicate a more dynamic view of the future (Hammerness, 2001; Senge, 2006). After all, it should not be surprising if teachers’ talk about mathematics instruction is “out ahead” of their enactments. Sfard (2007) suggested that “[w]e need a discursive change to become aware of new possibilities and arrive at a new vision of things. We thus often need a change in how we talk before we can experience a change in what we see” (p. 575).

Pursuing a different, but related, construct, Sherin (2001) adapted and applied Goodwin’s (1994) idea of *professional vision* to her work in documenting the evolution of one mathematics teacher’s perspective of classroom events. Over the course of her four-year collaboration with Mr. Louis, Sherin documented how his interpretation of classroom events captured on video from his classroom changed from a focus on his own pedagogical actions (i.e., what he should have done differently) to a focus on student ideas and the nature of mathematical discussions (i.e., accounting for what had actually transpired in classroom events). Over time, she documented which aspects of the classroom the teacher emphasized as being important with respect to mathematics instruction and learning, and the rationale behind his choices.

Though they are certainly related, each of these three ideas—Sherin’s professional vision, Hammerness’s (2001) teacher vision, and the vision of high-quality mathematics instruction (VHQMI) described in this paper—has unique characteristics that distinguish it from the others. Hammerness argued for attending to a teacher’s vision of instruction, whereas Sherin described
what could be considered Mr. Louis’s vision in instruction\(^1\) (or at least in video-observations of instruction). While the work reported here is in the vein of the former, in the world of research on teaching, where the bulk of writing describes end-goals or “best practices,” studies such as Sherin’s are helpful in that they provide some indication of what teachers’ learning paths might look like in coming to enact new forms of practice. In particular, she documented how a teacher began to focus more on the aspects of classroom activity to which researchers rather than teachers typically attend, and presented the case of Mr. Louis as “one possible trajectory in the development of teachers’ professional vision of classroom events” (p. 91). It is in this sense that the trajectories of perceptions of high-quality mathematics instruction presented in this report were developed, and in this sense that they differ from Hammerness’s (2001) vision constellations. That is, the VHQMI rubrics presented later model the development (and not merely constellations) of teachers’ and others’ visions of high-quality mathematics instruction as they progress toward a particular instructional vision drawn from mathematics education research (and described in the next section).

**Theoretical Perspectives**

In this section I describe an agenda for school mathematics and a theoretical perspective on classroom learning from which the rubrics described in this paper were developed.

**An Agenda for Mathematics Learning**

Following Carpenter and Lehrer (1999), I take learning mathematics with understanding—which I view as requiring the development of multiple strands of mathematical proficiency (Kilpatrick, Swafford, & Findell, 2001) and forms of mathematical practice (Common Core State Standards Initiative, 2010)—to be the overarching goal for what we hope

\(^1\) Thanks to Eve Manz for naming this distinction.
students will achieve in classrooms. The authors argued that this kind of learning is generative. “When students acquire knowledge with understanding, they can apply that knowledge to learn new topics and solve new and unfamiliar problems,” rather than perceiving each mathematical idea as representative of an isolated skill (p. 19). Based on evidence from their own classroom research, the authors proposed that mathematical understanding emerges from five forms of activity: 1) constructing relationships between current and more sophisticated ways of thinking; 2) building rich, integrated knowledge structures; 3) reflecting; 4) articulating what one knows; and 5) making mathematical knowledge one’s own by coming to perceive it as evolving and authoring one’s own learning. Concerning the last of these activities, Lampert (1990) identified students' development of intellectual authority as the key distinguishing feature between mathematics as a disciplinary practice and typical school mathematical activity. She suggested that, like mathematicians, students should engage in public analyses of the assumptions they make to generate their answers and the legitimacy of those assumptions. Gravemeijer (2004) further defined such intellectual autonomy of students as being able “to only accept new mathematical knowledge of which they can judge the validity themselves” (p. 109).

**Perspective on Classroom Mathematics Learning**

Underlying the agenda described above is a theoretical perspective on learning that is well-captured by Cobb and Bauersfeld’s (1995) “emergent perspective,” which involves the coordination of social and psychological perspectives on classroom activity. These researchers came to view social norms not as something defined and imposed solely by the teacher, but as patterns of activity co-constructed by teachers and students. In relation to the psychological aspect of their emergent perspective, they argued, “in making these contributions [to aspects of the classroom microculture], students reorganize their individual beliefs about their own role,
others’ roles, and the general nature of mathematical activity” (Cobb & Yackel, 1996, p. 178).

The authors suggested that neither the social nor the psychological perspective plays a more prominent role than the other, but that they are reflexively related and co-dependent. Carpenter and Lehrer’s (1999) five forms of mental activity align with an emergent perspective of classroom learning, in that they imply both individual knowledge construction and mathematics learning as a social practice.

But to understand how such learning opportunities can actually be achieved requires an explication of the roles played by teachers and students mutually engaged in building classroom culture and microculture, for which Rogoff, Matusov, and White’s (1996) “community of learners” instructional model is helpful. These authors described two models of instruction—adult-run and children-run—as being not on opposite ends of a continuum, but actually quite closely related in that both correspond to a view of learning as a one-sided process, either managed by an expert transmitter of knowledge, or driven by active acquisitionists of knowledge. Their alternative model recast students’ and teachers’ roles as being mutually engaged in a shared endeavor, where students learn “as they collaborate with other children and adults in carrying out activities with purposes connected explicitly with the history and current practices of the [classroom] community” (p. 390).

Rogoff and colleagues’ (1996) argument is not that students will fail to learn mathematics in adult- or children-run instruction, but that across different instructional models students will “learn a different relation to the subject matter and to the community in which the information is regarded as important” (p. 391, italics added for emphasis). This helps clarify the roles one would expect to find teachers and students playing as they co-construct patterns of activity in the
classroom (Cobb & Yackel, 1996): the teacher as a more knowledgeable partner, responsible for ensuring that classroom mathematical practices come to resemble those of the discipline.

**Defining “High-Quality Mathematics Instruction”**

This perspective on mathematics learning informed the vision of high-quality mathematics instruction I propose in this article. It is defined in terms of three related dimensions of classroom instruction—role of the teacher, classroom discourse, and mathematical tasks.

**Role of the teacher.** Research conducted in classrooms in which teachers are supporting students in learning mathematics with understanding has, collectively, characterized the teacher as a “more knowledgeable other” (Rogoff et al., 1996)—one who proactively supports students’ learning by intentionally structuring classroom activities to afford genuine mathematical work and then co-participating in that work. Though it is non-content-specific, Engle and Conant’s (2002) description of well-designed learning environments provides a succinct account of what might best capture instructional visions of the teacher’s role at the highest level of sophistication. Drawing, in part, on the work of mathematics education researchers, they argued that environments should support problematizing ideas (Hiebert et al., 1996), give students authority (Lampert, 1990), hold students accountable to others and to shared disciplinary norms (Cobb, Gravemeijer, Yackel, McClain, & Whitenack, 1997), and provide students with relevant resources.

In keeping with the co-participation notion, those who take Engle and Conant’s view might suggest that the teacher should (a) engage *with* students in mathematical argument (Lampert, 1990); (b) play a proactive role in supporting and scaffolding students’ talk (Fraivillig, Murphy, & Fuson, 1999) by utilizing students’ explanations and questions as lesson content (Lappan, 1993), choosing appropriate moments to share essential information such as
conventions and alternative methods (Hiebert et al., 1997), or articulating important ideas in students’ methods (Cobb, Boufi, McClain, & Whitenack, 1997); or (c) ensure that “the responsibility for determining the validity of ideas resides with the classroom community” and not solely with the teacher or textbook (Simon, 1994, p. 79), which requires the teacher to consistently keep students positioned as the thinkers and decision-makers (Staples, 2007) while they move toward the mathematical horizon of important ideas in the discipline (Ball, 1993).

Generally, researchers have identified a three-phase classroom activity structure as one in which teachers can more easily and effectively play such a role (Van de Walle, Karp, & Bay-Williams, 2012)—termed launch-explore-summarize by some curriculum developers (Lappan, Fey, Fitzgerald, Friel, & Phillips, 1998). In well-executed lessons of this kind, the teacher poses a problem and ensures that all students understand the context and expectations (Jackson, Garrison, Wilson, Gibbons, & Shahan, 2013), students develop strategies and solutions (typically in collaboration with each other), and, through reflection and sharing, the teacher and students work together to explicate the mathematical concepts underlying the lesson’s problem (Stein, Engle, Smith, & Hughes, 2008; Stigler & Hiebert, 1999). As is clear in this brief summary of the literature, much of the teacher’s work is accomplished through classroom discourse.

Classroom discourse. In general, I take the development of a discourse community (Hufferd-Ackles, Fuson, & Sherin, 2004; Lampert, 1990) to be the goal for talk in any mathematics classroom. In such classrooms, whole-class conversations—which, as noted above, presumably follow students’ collaborative efforts to solve problems—include student-to-student talk that is student-initiated and does not depend entirely on the teacher for soliciting strategies, pressing for explanations, or explicating connections between mathematical ideas. Fundamental to the establishment and success of a discourse community is the nature of classroom talk (what
Thompson, Philipp, Thompson, and Boyd’s (1994) distinction between conceptual and calculational orientations to explanations of solutions is helpful in this regard. Calculational orientations treat problem solving as the production of correct, numerical answers and articulations of strategies as step-by-step accounts of procedures followed. But the development of a discourse community requires conceptual orientations, which promote a “rich conception of situations, ideas, and relationships among ideas” (Thompson et al., 1994, p. 86).

Within classroom discourse communities, (a) talk that is of a conceptual orientation is likely comprised of mathematical arguments, relations among multiple strategies, and explorations of errors and contradictions (Kazemi & Stipek, 2001; Simon & Schifter, 1991); (b) teachers’ questions drive investigations, help students explain their problem-solving strategies, and help the teacher understand students’ thinking (Borko, 2004); (c) student discussions spawn new questions and help students recognize confusions (Engle & Conant, 2002); and (d) students hold others accountable for producing sufficient mathematical arguments by repeating their own or others’ questions until satisfied with answers (Hufferd-Ackles et al., 2004). Of course, discourse communities are unlikely to develop unless classroom participants have opportunities to engage in rich mathematical work, typically initiated by the tasks that are posed.

Mathematical tasks. Hiebert et al. (1997) described four characteristics of high-quality mathematical tasks. They suggested that tasks should allow students to treat situations as problematic (i.e., as something to think about rather than a prescription to follow); make the mathematics the intriguing or perplexing part of the situation (rather than merely the context); offer students chances to use skills and knowledge they already possess; and be suitable to the available tools. Tasks of this nature, they argued, support students in developing problem-solving strategies and allow for "insights into the structure of mathematics" (p.23).
Members of the 1990s QUASAR project (Silver & Stein, 1996) characterized the quality of tasks that mathematics teachers choose to use in terms of the tasks’ potential for engaging students in cognitively complex work. Stein and colleagues developed a four-level rubric for analyzing the cognitive demand of any task (Stein, Grover, & Henningsen, 1996; Smith & Stein, 1998). All of the above descriptors align with Stein and colleagues’ characterization of the highest level of tasks, which they define as having potential to engage students in “doing mathematics”—solving challenging, ambiguously defined problems without the suggestion of a particular procedure or path to a solution.

Summary. The vision of high-quality mathematics instruction just described represents both what the MIST research team and leaders in each partnering district took as the overarching goal for mathematics teaching and learning, and what I took to be the most sophisticated instructional vision in modeling trajectories of their development. Thus, having articulated its theoretical underpinnings and orientation, I describe the development and application of the VHQMI instrument in the next sections. The first provides a description of the process by which the instrument’s rubrics were developed using the first two years of data of the MIST project. The second section documents the coding of interview transcripts from years one through four and briefly illustrates the kinds of findings such an analysis yields.

Development of the VHQMI Rubrics

As previously stated, the rubrics model of trajectories of instructional visions along what have been identified in the literature (and by study participants) as critical dimensions of mathematics classroom teaching and learning. To build these models, I relied on the few previous attempts in the literature to identify such critical dimensions as well as interview data collected in the first two years of the MIST project. In addition to the most sophisticated
instructional vision already described, previous research provided categories for an initial coding scheme and, in a very few cases, some empirically based accounts of teachers’ evolving practices or visions along a particular dimension. Again, the goal of my analysis was to fill in what I perceived as missing pieces in the literature: trajectories of less to more sophisticated instructional visions that teachers and others articulate “on the way” to the sophisticated vision that is described in the literature. In the next several paragraphs, after describing the data that informed the development of the VHQM1 rubrics, I explain my methods for integrating an analysis of those data with previous findings concerning critical dimensions of mathematics classroom instruction and learning, which resulted in the four rubrics included in the appendix.

**Data sources and collection.** At the outset of the MIST project, in each district, 6-10 middle schools were identified as representative\(^2\) of the range of student demographics and school success (i.e., in student achievement) in the district (30 schools total). Throughout the four years of the study, the participating schools remained constant. The roster of participants, however, evolved as individuals changed schools or roles, or, in a very small number of cases, chose to terminate their participation in the study. Since the larger project’s focus was at the district level, changes in study participants resulting from fluctuation in school and district personnel was viewed not as attrition, but as a natural phenomenon occurring in any large school district. So that we could maintain a representative sample, we enlisted new participants each year, keeping the number of participants (approximately 50-60 per district) fairly consistent across all four years, but also linking data within each participant regardless of the number of years of participation.

\(^2\) Schools were sampled purposively to include a range of school performance levels, ethnicity and other relevant factors to adequately represent each district as a whole.
A variety of data were collected for the MIST project (e.g., video-recorded observations of teaching, surveys, assessments of mathematical knowledge for teaching, artifacts), including annual interviews with district leaders (e.g., chief academic officers; superintendents; mathematics supervisors; directors of district departments such as curriculum and instruction, assessment, leadership, and special education), and with mathematics teachers, principals, and mathematics coaches in each participating school. The interviews, which ranged from 45 to 90 minutes in length, were audio-recorded and later transcribed. As is the case with semistructured interviews (Brenner, 2006), they followed a set of guiding questions (that were customized for each district), but were conducted flexibly, intended to afford opportunities for conversations rather than a rigid sequence of questions. The interview addressed topics such as participants’ responsibilities in their school or district, the nature of collegial interactions, plans for improvement, perspectives on current initiatives, supports and resources provided, and accountability relationships. The VHQMI-related portion of the interviews began by asking participants the following question: “If you were asked to observe a teacher's math classroom for one or more lessons, what would you look for to decide whether the mathematics instruction is high quality?” Depending on the participant’s response, we asked, “Why do you think it is important to use/do _____ in a math classroom? Is there anything else you would look for? If so, what? Why?”

The purpose of this particular question was twofold. First, with teachers, it was an attempt to circumvent the “say-do” problem (Gougen & Linde, 1993), a well-known obstacle in the social sciences in that self-reporting typically does not yield reliable data concerning participants' own practices. We asked our participants to imagine and talk about the activities in a classroom of some hypothetical other in an attempt to release them from some of the pressures
of accurately describing their own classroom and practices, and tendencies to foreground more socially desirable aspects of instruction to the exclusion of those perceived as less desirable. Secondly, and more importantly, in asking participants to place themselves in the role of observer, we hoped to ascertain the lens with which they would view a mathematics classroom (Sherin, 2001). That is, we could interpret their responses to mean, “this is what matters in a mathematics classroom”—the aspects of the classroom on which they focus to determine the quality of instruction. This would enable us to establish the kinds of things they might attend to when observing a mathematics classroom (e.g., what the teacher or students are doing, the nature of mathematical tasks or classroom discourse), and also assess the sophistication of their criteria.

Based on analyses of Year 1 interviews (described below), we decided to prompt specifically in subsequent years on three topics drawn from the literature that seemed especially prevalent in participants’ responses, and therefore fruitful for assessing participants’ instructional visions: 1) the role of the teacher; 2) the role of classroom discourse; and 3) the nature of mathematical tasks used with students. Thus, for each of these three topics the participants did not identify spontaneously, we prompted by asking, respectively, 1) What are some of the things that the teacher should actually be doing in the classroom for instruction to be of high quality? 2) Can you please describe what classroom discussion would look and sound like if instruction was of high quality? 3) What type of tasks do you think the teacher should be using for instruction to be of high quality?

---

3 It is worth noting that, additionally, since we were also collecting video data in each teacher’s classroom, we did not need teachers to report on their own classrooms and instructional practices.
Sampling. For the purpose of developing the VHQMI rubrics, I examined transcripts from 38% of the 177 interviews conducted with teachers, principals, and mathematics coaches in Year 1 (January, 2008). This included interviews conducted with every participant at each of ten schools (at least two from each district), which were sampled to provide wide variation in the ways participants talked about mathematics instruction, as indicated in case summaries written for each school. (The exclusion of district leader interviews was not purposeful, merely a consequence of using a school-based sampling plan.) In an attempt to capitalize on potential changes in individuals’ ways of describing high-quality instruction, I also analyzed transcripts from interviews conducted in Year 2 with any participant whose Year 1 transcripts I had analyzed, totaling 23% of the 184 Year 2 (January, 2009) interviews conducted with teachers, principals, and coaches. Sample sizes are listed in Table 1.

Table 1

<table>
<thead>
<tr>
<th>District</th>
<th>Schools</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Principals &amp; asst. principals</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Teachers</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Math coaches</th>
<th>Year 1</th>
<th>Year 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>12</td>
<td>30</td>
<td>31</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>34</td>
<td>35</td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>33</td>
<td>32</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>34</td>
<td>35</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sampled for VHQMI development

<table>
<thead>
<tr>
<th>District</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 1</th>
<th>Year 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(of 30)</td>
<td>(of 30)</td>
<td>(of 34)</td>
<td>(of 38)</td>
<td>(of 131)</td>
<td>(of 129)</td>
<td>(of 12)</td>
<td>(of 17)</td>
</tr>
</tbody>
</table>

Analysis. In the first phase of my analysis, I examined Year 1 data to identify dimensions of classroom instruction that would adequately characterize the nature of the participants’ responses (and around which I would then develop the VHQMI instrument). Focusing in particular on the portion of the interviews including the question and probes mentioned above, I collected approximately 240 statements (each of which was one or more sentences) from Year 1 interviews with principals, teachers, and coaches. By “statement” I am referring to what has been characterized as a meaning unit: “words, sentences or paragraphs containing aspects related to
each other through their content and context” (Graneheim & Lundman, 2004, p. 106)—a flexible, but meaningful unit of analysis when using complex interview data (Campbell, Quincy, Osserman, & Pedersen, 2013).

My initial classification of those statements was guided by a provisional list of codes (Miles & Huberman, 1994) drawn from three attempts in the literature to glean a coherent set of crucial dimensions of mathematics classroom instruction. In Figure 1, I map the critical dimensions of mathematics classroom instruction identified by Carpenter and Lehrer (1999) and by Franke, Kazemi, and Battey (2007) onto the dimensions described by Hiebert et al. (1997). In each case, the authors attempted to describe a set of features of mathematics classroom instruction they viewed as critical for providing opportunities to learn mathematics with understanding—the dimensions that “matter.”

Combining all three sources, the fourth column of Figure 1 represents the provisional coding categories I used at the outset of my analysis.

These six categories accounted for approximately 60% of the statements that I coded, although very few responses pertained to social culture (2%), mathematical tools (1%), or equity (3%). Following Strauss & Corbin’s (1998) open coding technique, I then classified statements

4 The ‘gaps’ in these lists should not be interpreted as omissions; they are typically a consequence of arrangement and classification choices. After all, all of the authors acknowledged systemic relationships among their dimensions.

5 That the initial dimensions were a good fit for only 60% of the responses sampled does not necessarily indicate a shortcoming of the abovementioned researchers’ frameworks. Again, their goal was not to classify teachers’ visions for mathematics instruction, but to describe high-quality practice itself—those aspects of the mathematics classroom they viewed as essential for supporting children in learning with understanding.
for which the provisional categories did not account into categories based on shared properties, such as whose behavior the statement pertained to (e.g., teacher, students, or both) or which aspect of the learning environment was emphasized (e.g., nature of classroom tasks, structure of lessons, etc.). This resulted in three candidates for additional categories: student engagement in classroom activity; assessing student thinking; and lesson structure (i.e., important components of a lesson, or the “flow” of a lesson). However, only “student engagement,” which was the most frequently named aspect of the classroom for which participants would look (29% of the statements), was strongly populated compared with the provisional categories, with only 5% of statement pertaining to lesson structure and 2% to assessing student thinking.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Role of the Teacher</td>
<td>Building relationships for doing and learning mathematics</td>
<td>The Role of the Teacher</td>
<td></td>
</tr>
<tr>
<td>Equity and Accessibility</td>
<td>Equity and Accessibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nature of Classroom Tasks</td>
<td>Tasks or activities students engage in and the problems they solve</td>
<td>Supporting discourse for doing and learning mathematics</td>
<td>The Nature of Classroom Tasks</td>
</tr>
<tr>
<td>Social Culture of the Classroom</td>
<td>Classroom normative practices</td>
<td>Establishing norms for doing and learning mathematics</td>
<td>Social Culture and Norms</td>
</tr>
<tr>
<td>Mathematical Tools as Learning Supports</td>
<td>Tools that represent mathematical ideas and problem situations</td>
<td>Mathematical Tools as Learning Supports</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1. Summary of central aspects of mathematics instruction identified in three works.*

For purposes of instrument development, my goal was to be able to adequately characterize VHQM with as few dimensions as possible. Therefore, based on the distributions I
observed in the first phase of my analysis (of Year 1 data), I kept four of the nine potential
dimensions listed above in the final coding scheme: 1) role of the teacher; 2) classroom
discourse; 3) the nature of classroom tasks; and 4) student engagement in classroom activity. As
explained above, starting in year two, interview protocols were expanded to include specific
prompts associated with the first three. (We opted not to prompt for participants’ descriptions of
student engagement because we assumed that, from respondents’ perspectives, asking, “In what
would you expect students to be engaged?” would likely be similar to prompts regarding tasks
and discussion.)

The focus of the second phase of my analysis was to elaborate trajectories for each of
these dimensions. To do so, I waited until after the second year of the study so that I could use
interview data from the same individuals across at least two years. In addition to the
approximately 240 Year 1 statements mentioned above, I collected approximately 190 statements
from Year 2 transcripts. Similar to the process described by Fennema et al. (1996), I analyzed
responses within each dimension to identify patterns indicating potentially important qualitative
distinctions that could help elaborate developmental trajectories. My intent was to identify and
name levels comprising the trajectories, and illustrate those levels with examples (directly
quoting participants when applicable). Informing the directionality of the trajectories I was
attempting to model were descriptions of high-quality teaching from mathematics classroom

6 Although equity was not a dimension about which participants frequently commented in
response to the primary VHQMII prompt, responses to other portions of the interview did provide
some indication of participants’ views on this dimension. Other researchers from the MIST
project have developed a scheme for coding comments about equity and accessibility (cf.
Jackson & Gibbons, 2014), which, in the future, could provide a complement to VHQMII rubrics.
research. Specifically, I took the literature-based instructional vision previously summarized as representing the top levels of the rubrics, and, within each dimension, interpreted each interview response against that benchmark. In a few instances, participants’ responses provided quotable examples of these most sophisticated levels. More often, however, I used the interview data to identify criteria for making distinctions along the way to those points. Though far less often, previous research was helpful in articulating these lower levels as well; occasionally, participants’ responses matched authors’ descriptions of what is not high-quality instruction.

In many instances, research that has identified important variations in form and function relationships within mathematics instructional reform efforts (Saxe, Gearhart, Franke, Howard, & Crockett, 1999; Spillane, 2000) proved useful in differentiating between less and more sophisticated responses and contributed to building levels within each trajectory. In examining primary and secondary mathematics teachers’ classroom assessment strategies, Saxe et al. (1999) accounted for changes in teachers’ practices with respect to three sources of press. These included presses emerging at an institutional level (e.g., standards, curriculum, testing, professional development programs), a local level (e.g., parents, principals, colleagues, students) and an internal level (e.g. reflection on one’s practices with respect to personal values). As part of their account of the influences of these sources of press, the authors documented shifts over time in teachers’ reported uses of assessment forms such as open-ended tasks and scoring rubrics, but reported that many teachers used new forms to serve old functions. For example, a majority of participating teachers reported an increased use of open-ended assessment items—at least once per week. However, half of the teachers reported using an accompanying rubric for interpreting and using student responses much less frequently; many used it as little as once or twice per year. As a consequence, the authors argued, many teachers were using the new, open-
ended form of assessment to serve the familiar function of evaluating answers as right or wrong, rather than using the tasks to diagnose students’ thinking.

Spillane (2000) extended the form-function distinction to district policymakers’ conceptions of aspects of mathematics instruction reform, three of which they often described as “problem-solving,” “mathematical connections,” and “hands-on.” Spillane suggested, “[f]orm-focused understandings refer to pedagogical forms including learning activities, students’ work, instructional materials, and grouping arrangements. Functional understandings center on what counts as mathematical knowledge, doing mathematics, and learning and knowing mathematics” (p. 154). Spillane identified a pattern in district policymakers’ interpretations of reform initiatives, arguing that they often focused on the forms of reform, missing their intended functions as supports for student learning. For example, although a new emphasis on problem solving marked a change in district leaders’ agendas for classroom instruction, “it did not involve fundamental change in what counted as mathematics in the K through 12 curriculum” (p. 155). Likewise, district leaders described the need for “real-world connections” in mathematics instruction in terms of making mathematics more relevant and engaging for students but did not describe their function as using meaningful contexts to promote children’s sense-making and mathematical reasoning. Similarly, Spillane argued that leaders’ emphasis on “hands-on” tasks or “use of manipulatives” were still targeted at helping students learn procedures for solving tasks rather than providing tools to assist them in constructing new understandings.

In my analysis of participants’ responses that emphasized elements of reform mathematics instruction (e.g., hands-on tasks, use of manipulatives, formative assessments, high-level questioning), I differentiated between descriptions that included both new forms and thorough descriptions of their functions in terms of supporting students’ learning, and those that
promoted new forms but failed to indicate a shift away from old functions. For example, participants might stress the importance of students’ questions, stating that it is important to get all students participating and engaged, rather than suggesting that students’ questions be used to drive instruction. While both of these rationales suggest a function underlying the form of students’ questions, only the latter is specifically related to mathematical learning opportunities. My argument for making such distinctions is that an ability to articulate a strong rationale for employing aspects of mathematics instructional reform indicates a more sophisticated vision of high-quality mathematics instruction.

The form-function distinction applied to a majority of the trajectories. However, other differentiating qualities were more dimension-specific. In the results section, accompanying the description of each of the four VHQMI rubrics, is a review of such dimension-specific qualities identified in the literature that contributed to identifying meaningful levels.

The work of developing the rubrics spanned two years. Throughout the process, I regularly solicited feedback from the MIST project team, including multiple rounds of detailed, written reactions to each rubric from one of the project’s principal investigators. As a last step in developing the rubrics, I convened a task force of veteran project members with considerable experience in developing and applying qualitative coding schemes in making final refinements to the rubrics. Those colleagues first provided their written reactions to each rubric, raising questions about the appropriateness of leveling (e.g., whether two should be collapsed, or another split in two) and identifying imprecise wording. This list of potential areas of refinement focused our next step, which was independent coding of interview transcripts (i.e., identifying the appropriate level for each dimension to which a participant’s responses pertained), followed by discussion of how we interpreted and used the rubrics to assign scores. For this work, I
selected transcripts from all role groups (teacher, coach, principal, and district leader) that, based on my initial read, would collectively provide opportunities to employ all of the rubrics and at least a majority of levels in each. One interview at a time, the task force engaged in iterative rounds of independent coding and follow-up discussions. Through this process, we first came to an agreement on the appropriate number of levels in each rubric, and then identified ways in which definitions and examples of levels could be improved in order to clarify when a given code should be applied. These analyses yielded the rubrics presented in the appendix and described below.

**The VHQM1 Rubrics**

Although necessary for assessing the complexity of individuals’ instructional visions, the rubrics’ density makes their presentation in a journal article challenging. Therefore, in the paragraphs below, I introduce each of the four rubrics by providing an orientation to their composition and intent, describing their structural aspects, summarizing contributions from previous research, and characterizing the main themes present in distinguishing levels. Understanding some details, however, will likely require studying the rubrics themselves (the full versions of which are available upon request).

Each rubric consists of numbered levels. The highest (typically level 4) indicate the most sophisticated instructional visions, the literature-based definitions of which were described in the Theoretical Perspectives section. Here, I describe the levels leading up to those most sophisticated instructional visions. Grounded in our data, the literature, or both, these levels mark what I determined to be important qualitative distinctions that help to characterize individuals’ instructional visions (though, for the sake of simplicity, I use the shorthand language of “coding” or “scoring”). In the rubrics, descriptions of particular codes at each level are accompanied by
examples collected from either MIST data or, less frequently, from previous studies (the former indicated with direct quotes or paraphrases, the latter with listed references). Embedded in the narrative below are cases that exemplify progressions through the levels in order to illustrate how the rubrics are intended as models of trajectories of developing instructional visions over time. **Role of the teacher.** Even within theoretically compatible descriptions of the role of the mathematics teacher, researchers have emphasized different sub-dimensions. For example, Hiebert et al. (1997) described two major forms of guidance the teacher must provide, one concerning mathematical activity and the other classroom culture. Others have described the teacher’s role with respect to discourse (Staples, 2007; Borko, 2004), student thinking (Franke, Carpenter, Levi, & Fennema, 2001), and choosing and employing tasks and activity structures (Clarke, 1997). I identified three dominant ways that our participants characterized the role of the teacher: (a) conception of typical activity structure (i.e., the teacher’s general mode of instruction and role in classroom activities), (b) attribution of mathematical authority (i.e., students’ roles with respect to the mathematics being learned), and (c) influencing classroom discourse. With the exception of the lowest level (0), each level in the role of the teacher rubric is defined according to each of these three potential characterizations.

The lowest level (level 0, teacher as **motivator**) of the rubric pertains to responses that are limited to asserting that the teacher must be energetic and captivating (so that students will be sufficiently motivated to learn), with no mention of what the teacher should do with respect to content (which precludes defining this level according to the three characterizations noted above). Such comments are often accompanied by descriptions of high-quality instruction as an inherent trait (e.g., “Some people are just naturally very good at teaching”). A level 1 code requires that the participant specifies that the role of the teacher is to teach mathematics.
Specifically, at level 1 (teacher as *deliverer of knowledge*), the participant’s description suggests that the teacher has mathematical knowledge that must be imparted unto students, which is most effectively accomplished through efficiently structured lessons (in terms of coverage) in which the teacher directly teaches how to solve problems and provides time for practice.

At level 2 (teacher as *monitor*), descriptions suggest that students should play an active role in working together on mathematical tasks and that affording time to students for figuring out (or, more likely, reproducing) what the teacher has explained or demonstrated is important. Whereas at level 1 the image of the student’s role is one of receiving knowledge, level 2 responses describe students as playing a role in mediating what the teacher has explained, but with the teacher playing the role of “adjudicator of correctness” (Hiebert et al., 1997)—a new form serving an old function.

To earn a level 3 score (teacher as *facilitator*), the participant must describe the teacher’s role as facilitating student discovery during at least part of the lesson. At most, the imagined teacher introduces the day’s task and does the first part or two of that task with the class before turning it over to the students, then keeps students on the “right path” by asking questions. Or, the participant might suggest that the teacher’s role consists of very little because learning mathematics should be a process of “student-led” discovery. Level 3 responses often include reform-oriented rules to be followed in interaction with students (e.g., “answer questions with questions,” avoid “spoon-feeding”). Additionally, this level includes descriptions that emphasize posing problems and asking students to describe their strategies but do not describe engaging students in genuine mathematical inquiry (Kazemi & Stipek, 2001) through proactive, co-participation with students. For example, whereas at level 3 one might caution the teacher not to “tell” too much for fear of interrupting the “discovery” process, a level 4 vision more clearly
articulates the role of telling in terms of guiding the mathematics in meaningful ways (Lobato, Clarke, & Ellis, 2005; Chazan & Ball, 1995).

As examples of how individuals’ perceptions of the role of the teacher might develop through these levels over time, Figures 2 and 3 represent the progressions of two teachers, Anita and Kelsey, respectively, over four years. Anita’s Year 1 response was scored as a level 1, since it suggests that a teacher should teach some material and then provide time for students to practice doing it. (Had her description been limited to holding students’ attention, it would have been scored a level 0.) In Years 2 and 3, her descriptions included an emphasis on student-to-student mediation of the lesson that the teacher has just taught (level 2). In Year 4, her description shifted to suggest that the teacher should turn it over to the students, emphasizing introducing investigations rather than “giving a lesson” (level 3).

<table>
<thead>
<tr>
<th>Year</th>
<th>RT</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>“the teacher needs to have their attention of at least 90% of the students... You know, get up there, instruct, but don’t instruct the whole period. Allow kids to practice what they’ve learned... the guided practice is very important to assist them, everyday that you give a lesson.”</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>“I would look for students being able to explain to another kid, ‘Hey this is the reason why’... Accountable talk, I never even thought about it before because in school we didn't have partners and peers... But when you look at accountable talk and how it's really done, the kids are able to, they have to prove it... And so you know you kinda like, you come in and you lead 'em... [but] they're really, you know, having to prove it.”</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>“You can’t have the students doing everything by themselves, cause the teachers still needs to instruct, but she also still needs to walk around and facilitate and make sure that they're doing what they're supposed to do, but listening to them talking and speaking and seeing what their thinking is.”</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>“questioning strategies of the teacher... You’ve got to be prepared to ask those questions that’s gonna dig in deeper with the students... if a group is lost or if the student is lost, just, you know, help them get back on track. Again with the questioning thing... The students are given an investigation, ... the teacher kinda goes through and talks about it, introduces it and everything, ... The students are ... giving their different opinions on why it is or why it’s not. I mean I think that’s accountable talk; they’re able to justify their reasoning... And then they come together and they, you know, get everybody’s input and they decide on, you know, which way they want to go to solve the problem.”</td>
</tr>
</tbody>
</table>

Figure 2. Progression of Role of the Teacher (RT) scores for Anita, a district B teacher, over 4 years

Kelsey, whose interview excerpts are included in Figure 3, provided responses that were scored as level 3 for the first three years of the study because she consistently suggested that students should be provided opportunities (and be expected) to figure things out for themselves. Only in Year 4 did she suggest that the teacher should proactively orchestrate concluding whole-
class conversations by “taking notes on what people are saying and doing for the summary part” (level 4).

<table>
<thead>
<tr>
<th>Year</th>
<th>RT</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>“I would look for are the kids engaged in their groups or how the teacher is questioning them and then do like do the kids you know is there, are the kids being asked to do high level thinking and are they explaining their thinking and then does the teacher land the mathematics during the class period. You know in the summary do they, do they land the key math concepts with the kids”</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>“the kids are engaged in figuring out for themselves versus like this is how you do it kind of thing… [while the teacher is] going around asking kids questions and a helping them, either guiding their question or helping them think further or think about it a different way.”</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>“having [students] work in groups and letting them work on problems and solve ‘em and then discuss ‘em … and, you know, not giving away too much information to them.”</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>“the kind of questions [teachers] are asking when they’re walking around, … if the kids are required to think, and they’re not giving away everything to them [when] they go to their table. [The teacher would be] asking questions, walking around, even taking notes on what people are saying and doing for the summary part… you kinda like just don’t want to have the kids who have the perfect paper present, you wanna have the, like maybe some kids who got part of it present—a misconception that you saw at a lot of tables—you might wanna get that on the table. [And], kids who really landed their math, maybe things you heard that would be good examples of academic talk, quotes that you could give them so that they could learn to talk academically better with each other.”</td>
</tr>
</tbody>
</table>

**Figure 3.** Progression of Role of the Teacher (RT) scores for Kelsey, a district A teacher, over 4 years

**Classroom discourse.** Across the transcripts I analyzed that included statements about talk in the mathematics classroom, participants’ ways of characterizing its importance largely resembled those of research on classroom discourse. Some comments regarded *patterns and structure* of talk (“students should ask each other questions instead of asking the teacher”), some the general content or *nature* of talk (students should be “actually talking about math”); and others the presence or quality of specific components of talk, including *students questions*, *teacher questions*, and *student explanations*. Unlike the others, the classroom discourse rubric specifies levels for each of five sub-dimensions, and each is intended to be scored separately. Due to space constraints, included in the abbreviated rubrics in the Appendix and illustrated with examples of individuals’ progressions below are only the more general *patterns and structure of talk* and *nature of talk* sub-dimensions of the rubric.

As previously explained, Thompson et al.’s (1994) distinction between conceptual and calculational orientations to explanations of solutions proved very helpful in distinguishing
levels within each sub-dimension of the discourse rubric. Also helpful was the work of Hufferd-Ackles, Fuson, and Sherin (2004). Based on a year-long case study of one elementary teacher’s urban classroom, these researchers generated a developmental trajectory of the process by which the teacher and students came to establish a classroom community in which members influenced each other’s learning through meaningful mathematics discourse. Their framework described four levels, moving from teacher-directed instruction with minimal input from students to an environment in which the teacher is a co-learner and students are co-evaluators of other members’ work. While it certainly helps in articulating a sophisticated vision of high-quality classroom discourse, the authors’ proposed trajectory pertains to each school year’s progression toward a math-talk community within each classroom, and not the ongoing development of an instructional vision. Consequently, I could not simply adopt the model for interpreting our interview data. However, milestones within a (likely) more protracted development of an instructional vision are not completely dissimilar to those within an annual progression toward classroom community. Since multiple aspects of Hufferd-Ackles and colleagues’ model could potentially be included in an individual’s current vision of high-quality classroom discourse, their model was helpful in working to identify meaningful distinctions within our interview data.

Within the patterns and structure of talk sub-dimension, responses at the lowest level (level 1) suggest a one-way, teacher-to-student pattern of talk such as the initiation-response-evaluation (IRE) pattern described by Mehan (1979). Other levels of this code are used when the participant indicates that talk should not be limited to a teacher-to-student pattern (level 1) but should occur between students (level 2), whether that is in small groups only (level 2) or in
whole-class settings (level 3), and whether the teacher’s role in such conversations is central (level 3) or de-centralized (level 4).

Figure 4 provides a snapshot of one teacher’s progression in describing patterns and structure of talk in her instructional vision. In Year 1, Penny’s mention of talk was limited to a teacher-to-student pattern (level 1). In Year 2, she added an emphasis on group work (level 2), and then, in Year 3, suggested that students should be engaged in whole-class conversations after that group work (level 3). This last response was not scored at level 4 because the description places the teacher at the center of the conversation rather than inviting and orchestrating student-initiated discussion.

<table>
<thead>
<tr>
<th>Year</th>
<th>PS</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>“The modeling of what is being taught … Like today we’re going to use benchmarks to develop strategies to estimate fractions. When we talk about that hopefully, the children will get some idea of how you look at the numerator and the denominator by modeling, by using the number line… you’re able to show the students what you’re expecting them to do”</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>“the modeling and then the follow through with practice work that would involve group work because our goal is to have group instruction, which, you know, if the students are not able to get the instruction from the teacher then by working with their groups they could possibly either support their learning or else begin with the learning that they didn’t get [from] teacher instruction.”</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>“I would probably do groups of four if I were to do that because in order to write down your work, you need more than two people. So, and then, all of `em having an input into what is being presented on their paper… [and a whole class discussion] would be teacher led. And each person having a chance to say what they think about a particular topic, but it’s hard to have discussion because waiting your turn is not easy for these students. They wanna say it as soon as it comes to their mind, and listening to each other. They don’t, they don’t know how to do that.”</td>
</tr>
<tr>
<td>4</td>
<td>--</td>
<td>n/a (not in study Year 4)</td>
</tr>
</tbody>
</table>

Figure 4. Progression of Patterns & Structure of Talk (PS) scores for Penny, a district B teacher, over 3 years

Figure 5 represents another individual’s progression. At the outset of the study, Jacob was a teacher, but switched to the role of coach beginning in Year 3. In the first year, his interview response included no mention of the importance of talk, and was therefore not code-

7 The last of these should not be confused with the passive characterization of the role of the teacher in level 3 of the rubric—a de-centralized role does not mean passive; it refers to proactive co-participation.
able for patterns and structure of talk. In Year 2, his description included a clear emphasis on the importance of students talking to each other, but was limited to partner or small group work (level 2). In Year 3—the year he became a coach—his description of high-quality instruction began to emphasize the importance of whole-class “closure” conversations, where students present their work (level 3). In Year 4, his articulation of that concluding conversation describes students initiating turns of talk in response to each other (level 4). In relation to the notion of an instructional vision, Jacob’s last response is particularly interesting. He commented on the relative infrequency with which he was observing students talking to each other about their work in a whole-class conversation, but his description suggests that helping teachers achieve such patterns of talk was a goal in his work as a coach.

<table>
<thead>
<tr>
<th>Year</th>
<th>PS</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no mention of classroom talk</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>“just getting kids to feel comfortable talking about math… I think it would come out in classroom management, you know, walking around the classroom. When that kid asks you a question and you say ‘Well, did you ask your partner?’”</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>“I love when I see the closure where the teacher asks one or two questions and I see the kids doing the majority of the talking… they need to be presenting their work and those kinds of things. I don’t see a closure as being effective if it’s a 10-minute lecture on what you were supposed to have learned, because then if I’m a student in that class, then group work doesn’t mean as much because I’ll just wait until the teacher explains it to me.”</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>“A big piece for me is I want to see closure… not where the teacher says well today I expected you to learn blah, blah, blah and now I’m going to give you the answer, but pulling from the students and sharing the work… I have not seen it many times, but I really would love to get to the point where I’m seeing student and a student talking about each other’s work, and a lot of the observations I have made I’m seeing student to teacher and student wanting the teacher to validate their work and, and I don’t, and I know ideally, in an ideal setting you want to see the student to student discussing each other’s work and responding to each other because that means they’re listening and they’re really thinking about it.”</td>
</tr>
</tbody>
</table>

Figure 5. Progression of Patterns & Structure of Talk (PS) scores for Jacob, a district D teacher/coach, over 4 years

The *nature of classroom talk* code is used when the participant’s description of classroom talk specifies what that talk should be about or what it should consist of (e.g., questions, explanations, mathematical arguments). Although this code represents a general characterization of the kind of talk in the classroom and is not specific to teachers or students, it is applied only when the participant’s description suggests that students are involved in conversation with each
other and, therefore, does not have a level 1 application. The three levels distinguish between
descriptions that mention the content of students’ talk as something to consider but with little
specificity (level 2), those that provide a more detailed description of the content of classroom
talk, but with a calculational orientation (level 3), and those that stress that talk should be
conceptual in nature (level 4).

In Figure 6 are excerpts from interviews with Natasha, a teacher who was part of the
study in years two through four. Her progression in describing the nature of talk that one would
observe in high-quality instruction is tied to the notion of “accountable talk” (Michaels,
O’Connor, & Resnick, 2008), a focus of ongoing professional development in her district during
the years of the study. Across the three years, we observed an expansion of her conception of
that construct, shifting from an emphasis on staying on topic in Year 2 (level 2), to a more
detailed description of students arguing about answers in Year 3 (level 3), and finally, in Year 4,
a suggestion that in discussing their solutions, students should “compare and contrast how [their
solutions] are the same and different” (level 4). This last comment earned the highest score, since
it suggests more conceptually oriented talk than just discussing the correctness of calculations.

<table>
<thead>
<tr>
<th>Year</th>
<th>NT</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n/a (not in study year 1)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>“[The students are] doing what we call accountable talk, which is they’re talking about the objective or the goal… they can’t talk about their boyfriends, their girlfriends, what’s going on, no fighting, this or that. They can only talk about what we are discussing in math.”</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>“It’s accountable talk in groups…When they’re sitting in the groups, sometimes they will argue about whether the answer is wrong or right, and they will really press each other about the answers.”</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>“[Students] should be using what we call accountable talk… like today, one student was arguing with another one about the answer. And she’s like ‘Well, how did you get that answer?’ And they argued about how the answer was. Well, one of the students said, “Well, your answer’s right because this step was wrong.” And they said, “Well, how did you do it? Explain does this work? How does it apply? How did you use to find this? Compare and contrast how these two, mine and yours, are the same and different.””</td>
</tr>
</tbody>
</table>

Figure 6. Progression of Nature of Talk (NT) scores for Natasha, a district B teacher, over 3 years

The expanded version of the discourse rubric includes three additional codes: student
questions, teacher questions, and student explanations. These should be interpreted as ways of
specifying the *nature of classroom talk* code. Because they pertain to specific aspects of talk, they are typically used less often than the broader two codes described above. Their application is limited to instances when the participant emphasizes the importance or presence of such things in particular (e.g., “students should be asking questions”; “I would listen to the teacher’s questions”; "The teacher would be asking questions in order to...") or characterizes their nature (often by providing an example) (e.g., "The students would be asking questions like, 'Will that work in all cases?'”; “the teacher would be asking high-order questions”; “students would be asking each other how they did the problem”).

**Mathematical Tasks.** The rubric for coding participants’ responses pertaining to mathematical tasks builds on Stein and colleagues’ four-level rubric for analyzing the cognitive demand of any task (Stein, Grover, & Henningsen, 1996; Smith & Stein, 1998). At the lowest level of that rubric (level 1, “memorization”) are tasks that require only the reproduction of previously learned or demonstrated facts, rules, or formulae, with no problem solving. Levels 2 and 3 both pertain to tasks that require the use of procedures to solve problems, but are differentiated by whether the procedure that is called for or implied is closely connected to (level 3, “procedures with connections”) or isolated from (level 2, “procedures without connections”) the mathematical ideas that underlie the procedure. As noted previously in my definition of high-quality mathematics, the highest level tasks (level 4) have potential to engage students in “doing mathematics”—solving challenging, ambiguously defined problems without the suggestion of a particular procedure or path to a solution.

Boston (2012) adapted Smith and Stein’s (1998) levels of cognitive demand to create a rubric for assessing the potential cognitive demand of tasks in their inscribed form—one of multiple rubrics included in the Instructional Quality Assessment (IQA) Mathematics Toolkit. In
adapting Smith and Stein’s categories for the purposes of creating an IQA rubric, Boston kept levels 1 and 2 consistent, but made a slight modification to the higher-level categories. Whereas Smith and Stein emphasized the level of ambiguity of a task, Boston’s key distinction between levels 3 and 4 is that tasks of the highest quality require students to explain their reasoning behind a solution or conclusion (which could include both “doing mathematics” and “procedures with connections” tasks).

Smith and Stein’s (1998) rubric, along with Boston’s (2012) adapted version, provided a useful scheme for considering the quality of tasks that our participants described. However, the researchers’ categories were developed for the purpose of analyzing mathematical tasks themselves, and not educators’ visions of tasks. Consequently, their descriptions of tasks did not sufficiently resemble the ways that all of our participants talked about tasks. For example, none of our participants suggested that high-quality tasks require only that students reproduce memorized facts or definitions. In some cases, however, descriptions resembled “procedures without connections,” in that participants emphasized tasks that provide students with opportunity to practice a procedure before then applying it conceptually to a problem. As Hiebert et al. (1997) cautioned, tasks such as these promote the separation of conceptual understanding from procedures and contribute to students’ difficulty in adjusting procedures to solve different types of problems. Thus, the form-function distinction was useful in distinguishing levels of responses. Responses varied in the category (or form) of task emphasized as well as the reason for using such tasks (or function).

On the VHQMI tasks rubric, at the lowest level (level 0), responses suggest that the participant either (a) does not view tasks as inherently higher- or lower-quality, or (b) does not view tasks as a manipulable feature of classroom instruction (e.g., "We're supposed to be using
the CMP book, which is pretty much… what the teacher should say”). At level 1, responses suggest that tasks can vary in quality, but, as alluded to above, suggest that tasks should allow for procedural practice before application. To earn a 2, the response must describe the nature of tasks as being reform-oriented (e.g., hands-on, problem-solving), but either fail to describe a function or describe the function in terms of increasing interest levels and student engagement. Level 3 requires a more sophisticated description of high-quality tasks (e.g., should allow for multiple solutions, building on prior knowledge), which suggest that the function of such tasks is to support students’ conceptual understanding (and not merely achieving correct answers) or orientation to mathematics (i.e., to promote the notion that there are multiple ways of solving problems). Following Smith and Stein (1998), a level 4 description frames the rationale for high-quality tasks in terms of supporting students’ learning and “doing mathematics” (i.e., making and testing conjectures, examining whether a strategy will work in all cases, etc.). As an example, at a level 3, the potential for multiple solution paths is viewed as a desirable characteristic in and of itself, while at a level 4, multiple solution paths might be emphasized because they provide content for whole-class discussion and opportunities for making connections between ideas.

<table>
<thead>
<tr>
<th>Year</th>
<th>MT</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>“Problem solving—the kids are actually thinking about what they are doing, and not just doing rote”</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>“Problem solving, various kinds… they’re comparing, they’re contrasting, they’re applying what they’ve learned from one question to another, not just ‘here’s five problems that are all similar, now here’s five more.’”</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>“The kids are thinking… they’re actually problem solving… letting them discover on their own how to do the math maybe another way [rather than] ‘okay the steps are to do this’… One of the recent problems that I’ve done in here that I thought was a wonderful problem in CMP is when they’re learning how to solve linear equations and instead of just teaching them whatever you do to the left side of the equation you do to the right side of the equation, we’re balancing the equations—they have pictures of bags of coins and then single coins and they’re asking them how many bags are in a coin, and they kind of have to figure out how to balance out this equation without really having an x or a y or numbers, and I think that that lesson is really insightful for kids because… before the new math vocabulary comes out they actually figure it out for themselves and I think that that makes the rest of the unit so much easier. That to me is a high quality lesson—they’re understanding the math without me having to say ‘this is what you’re going to do today.’”</td>
</tr>
<tr>
<td>4</td>
<td>n/a (not in study year 4)</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 7. Progression of Mathematical Tasks (MT) scores for Doreen, a district A teacher, over 3 years*
Figure 7 provides an example of a teacher’s shift from level 2 to level 3. In the first two years, Doreen emphasized non-rote, problem-solving tasks, but in the third year expanded her description to emphasize tasks that provide opportunities for sense- and meaning-making.

A principal’s descriptions of mathematical tasks are included in Figure 8. In years one and two, Milt named multiple solution paths as an important characteristic of the problems that students are asked to solve, but did not specify a function for that form (level 3). In Year 3, his description linked the idea of multiple strategies to fruitful classroom discussion (level 4). (In Year 4, Milt became a district leader. In his interview he did not mention mathematical tasks as a dimension of the classroom he would consider, although the interviewer abbreviated the instructional vision section of the interview and excluded the prompt on mathematical tasks.)

<table>
<thead>
<tr>
<th>Year</th>
<th>MT</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>“So I’m looking for um teacher questioning and how they’re pushing the students to answer the question or help find the solution or multiple solutions. And having the expectations of finding multiple solutions… I like to listen to see how they solve problems and if teachers ask students to solve problems and to challenge each other about ‘Well, is there another way? Is that the right answer?’ and so on”</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>“what I'm looking for when I go in, is if the task has cognitive demands, if it is cognitively demanding… looking at how students are answering questions in multiple ways that they’ve had the answer, you know it helps us figure out where students are out, but it also helps us to look to see if the question is demanding”</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>“so I would look at the task and how that task was set up… we want students to think and, and you know there's multiple entry points so we want students to figure out problems in diff., come to the solution, but maybe get to that solution in different ways. We want them to challenge each other. We want them to you know agree or disagree. We want them to be, through socializing intelligence, to help each other figure out problems and solutions. We want them to be clear about what the task is. We want them to give each other feedback. We want, want the teacher to work, to be the facilitator of learning, not to stand and lecture and, and you know we want students to show evidence of why they came up with a solution they came up with, to defend their answer and even when they're challenged by the teacher, even though they may have the right answer, to stand, you know to try to stand by that answer and prove it, and so we use some prove it strategies with the students”</td>
</tr>
<tr>
<td>4</td>
<td>(no mention of mathematical tasks)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 8. Progression of Mathematical Tasks (MT) scores for Milt, a district A principal/district leader, over 4 years

**Student engagement in classroom activity.** As explained previously, “student engagement” was a frequent response to our question regarding high-quality mathematics instruction. However, when interviewers probed for elaboration, considerable variation in participants’ meanings emerged. To capture that variation, I initially attempted to employ Engle
and Conant’s (2002) delineation of three levels of student engagement. Although their guidelines for promoting and fostering disciplinary-specific engagement and learning were very helpful in developing a scheme for interpreting our participants’ descriptions of high-quality classroom activity, applications of the upper levels of that initial rubric were consistently redundant with those of the nature of talk dimension. Therefore, I retained only those levels that were necessary to capture the aspects of participants’ responses pertaining to “student engagement” that were not accounted for by other rubrics.

Responses at level one suggest a view of assessing the quality of instruction and learning by focusing on student behaviors or characteristics (and not solely the teacher's actions), but also suggest that the participant either takes for granted the quality of classroom activity (e.g., students should be “on task,” “doing whatever the teacher asked”), or privileges traditional forms of classroom activity (“students should be in their seats, listening, taking notes”). At level 2, descriptions espouse reform-oriented classroom activity (e.g., “Students should be up, moving around, using manipulatives”). Although such characterizations may fail to describe the content of students’ interactions (which would be coded on the discourse rubric), they suggest that the participant neither takes for granted the quality of the activity in which students are engaged nor views traditional forms of activity as high-quality.

Modeling Trajectories

The rubrics presented in this section (and included in the Appendix) provide a model of the trajectories that mathematics teachers’ and others’ instructional visions follow over time in settings in which leaders are promoting instructional practice aligned with mathematics education research. Though they are not discrete, but work together as a system (Hiebert et al., 1997; Carpenter & Lehrer, 1999; Franke, Kazemi & Battey, 2007), such a finite set of categories
can be useful for understanding how teachers and others view and talk about important aspects of mathematics instruction (Hiebert et al., 1997). Furthermore, my intention was to view participants’ accounts of important aspects of mathematics instruction in terms of their ongoing development (Simon & Tzur, 1999). However, I do not assume that the evolution of every individual’s instructional vision is linear or that the levels in the rubrics necessarily represent equal “increments of conceptual shift” (i.e., one should neither assume that the difference between 2 and 3 is the same across dimensions, nor that the difference between a 1 and 2 and a 2 and 3 is the same within a single dimension). Nor do I assume that all school personnel follow the same path or pace in that development. Rather, each rubric should be interpreted as a potential trajectory of developing instructional visions over time. Of course, these rubrics grew from analyses of data collected in districts whose leaders had particular goals for mathematics learning and teaching—a point I return to in the discussion section where I consider the applicability of the VHQMI instrument with personnel in any setting.

**Using the VHQMI Rubrics**

After completing the VHQMI rubrics, I oversaw the application of the rubrics in coding 932 transcripts of interviews conducted during the first four years of the MIST project. In this section I describe that process in order to demonstrate the feasibility of using the VHQMI instrument (including a report on reliability), and to provide some evidence of the validity of the data it generates—in this case, important information it provided to the MIST project.

**Coding**

Systematic VHQMl coding of Years 1–4 interviews occurred in two rounds. First, after the VHQMl rubrics were finalized by the task force mentioned above, coding of randomly ordered transcripts from 685 interviews conducted during the first three years of the MIST
project spanned October, 2010 to March, 2011. After the fourth year of data collection, an additional 247 interviews were coded in July, 2011.

As in the rubric development coding work explained previously, the VHQMI rubrics were applied to interview transcripts at the level of meaning units (Campbell et al., 2013; Graneheim & Lundman, 2004)—one or more sentences uttered by a participant to make a point about an aspect of teaching mathematics. Any statement made within the VHQMI portion of the interview could be coded on one or more dimensions. For each dimension or sub-dimension applied to a transcript, a final level was assigned based on the highest level coded. For example, in an entire interview transcript, multiple statements might be coded on the role of the teacher rubric at level 2, but if just one statement was coded at level 3, the final role of the teacher level for the participant’s interview would be a 3. Thus, the final assigned level can be interpreted as representing the greatest level of sophistication with which a participant was able to describe a particular aspect of mathematics instruction, though not necessarily with consistency. As explained in the following subsection, it was at the level of final scores that percent of exact agreement between coders was assessed, with the individual statement codes used as evidence in conversations about disagreements.

**Establishing reliability.** Prior to the initial round of coding, I provided a daylong orientation to using the VHQMI rubrics to members of the project team, which included multiple opportunities to independently code and then discuss results. Following that orientation, all interview transcripts (Years 1–3) were randomly ordered. I and one other member of the project team (referred to hereafter as the second-coders) came to consensus on the scoring of the first 20 transcripts on the randomized list, which served as the training reliability coding exercises for all potential coders. Potential coders began making their way through the list of transcripts, coding
1-3 at a time and discussing discrepancies with me. Coders were deemed reliable when, for at least the four most recently coded transcripts, they had achieved an overall rate of agreement of at least 80% and at least 75% agreement on each dimension or sub-dimension. After completing the training reliability phase, independent coding began (ranging from 3 to 20 interviews per week, depending on the coder), with weekly reliability checks performed by one of the two second-coders (alternating coders weekly). In the second round of coding (after year 4 data collection), I provided a refresher orientation to four of the coders from the previous round, after which they began coding with weekly reliability checks.

Table 2
Coder agreement rates

<table>
<thead>
<tr>
<th>Code</th>
<th>Years 1-3</th>
<th>Year 4</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role of the Teacher (RT)</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>Nature of classroom discourse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patterns &amp; structure of talk (PS)</td>
<td>0.80</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Nature of talk (NT)</td>
<td>0.74</td>
<td>0.76</td>
<td>0.75</td>
</tr>
<tr>
<td>Student questions (SQ)</td>
<td>0.89</td>
<td>0.93</td>
<td>0.90</td>
</tr>
<tr>
<td>Teacher questions (TQ)</td>
<td>0.77</td>
<td>0.73</td>
<td>0.75</td>
</tr>
<tr>
<td>Student explanation (SE)</td>
<td>0.87</td>
<td>0.83</td>
<td>0.86</td>
</tr>
<tr>
<td>Mathematical Tasks (MT)</td>
<td>0.80</td>
<td>0.74</td>
<td>0.78</td>
</tr>
<tr>
<td>Student engagement in classroom activity (CA)</td>
<td>0.85</td>
<td>0.86</td>
<td>0.85</td>
</tr>
<tr>
<td>Overall agreement rate</td>
<td><strong>0.81</strong></td>
<td><strong>0.79</strong></td>
<td><strong>0.80</strong></td>
</tr>
</tbody>
</table>

In both rounds of coding, each coder’s agreement rate was updated weekly for each dimension or sub-dimension and for all combined by calculating the rate of agreement between the coder’s scores and the final scores reached after the first and second coders had resolved any discrepancies. In the round of coding following Year 4 data collection, overall agreement on the teacher questions sub-dimension was only 0.69. Therefore, after clarifying the rules for scoring levels within that sub-dimension, coders re-examined all relevant transcripts, after which an additional 35 were randomly selected for double coding. The agreement rate among these additional double-checks was 0.77, bringing the overall rate for the teacher questions sub-dimension to 0.73.
discrepancies. A protocol was established for maintaining sufficient overall and dimension-specific rates of agreement (80% and 70%, respectively). For example, for any week that a coder’s rates dropped below the expected levels of agreement, the coder was asked to re-code the week’s transcripts, and additional checks were performed. In the first round of coding, 15% of the transcripts were randomly selected for double coding. In the second round, 19% were double-coded. Across all years combined, coders maintained an overall rate of agreement of 0.80.

Dimension-specific rates of agreement, averaged across all instances of double coding, are listed in Table 2, none of which was lower than 0.73.

**Coding Results**

Although a full report of the results of the coding analysis is beyond the scope of this paper, the following two examples provide an indication of the kinds of findings that such analyses yield. First, the analysis has allowed the project team to examine differences between districts, with respect to particular aspects of instructional practice. For example, as represented by Figure 9, average mathematical task (MT) scores for teachers in district A were higher than those of teachers in other districts—an unsurprising difference given that, when the study began, (a) district A teachers had, on average, at least five more years of teaching experience than teachers in other districts, and (b) district A had already been using a reform-oriented middle school textbook that emphasized complex tasks for several years, whereas other districts were first time adopters.
Second, employing the VHQMI instrument repeatedly over time provides insight into whether instructional visions are becoming more sophisticated. Given that the participants were in districts in which leaders were actively pursuing change in ways aligned with the vision of mathematics instruction on which the VHQMI rubrics are based, my expectation was that scores would increase over time. Of course, change (in any direction) at the school or district level could be attributed, in part, to personnel changes. Although the consequences of staffing policies and decisions is an important variable to examine, I was also interested in whether individual instructional visions change. Therefore, for each dimension or sub-dimension, I examined the average score for only those participants whose interviews were scored on that dimension in all four years (student questions and student explanations were omitted because they had only 1 and 3 observations, respectively).

The results, presented in Table 3, point to increased sophistication in participants’ articulations of each dimension of instruction over four years, including four dimensions for which average scores made at least one statistically significant increase. That the scores increase...
over time fits with the fact that data collection began as each district was initiating comprehensive reform efforts in middle school mathematics. It should be noted that increases in VHQMl scores between the first two years could be due, at least in part, to the addition of dimension-specific prompts to the interview protocol described above. But this does not account for increases in the following three years. Also, growth trends are not likely an artifact of coding since, as stated above, interviews from the first three years were randomly ordered so that coders’ work did not match the data collection chronologically.

Table 3  
Mean VHQMl scores by year for participants whose interviews were scored on that dimension every year

<table>
<thead>
<tr>
<th>Year</th>
<th>RT (0-4) [n = 44]</th>
<th>PS (1-4) [39]</th>
<th>NT (2-4) [23]</th>
<th>TQ (1-4) [14]</th>
<th>MT (0-4) [18]</th>
<th>CA (1-2) [12]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.98 (.88)</td>
<td>2.18 (.45)</td>
<td>2.83 (.58)</td>
<td>2.21 (.89)</td>
<td>2.39 (.50)</td>
<td>1.5 (.52)</td>
</tr>
<tr>
<td>2</td>
<td>2.48** (.95)</td>
<td>2.54* (.68)</td>
<td>3.00 (.67)</td>
<td>2.64 (1.01)</td>
<td>2.33 (.84)</td>
<td>1.42 (.51)</td>
</tr>
<tr>
<td>3</td>
<td>2.75# (.69)</td>
<td>2.72 (.76)</td>
<td>3.22 (.67)</td>
<td>2.43 (1.16)</td>
<td>2.83* (.79)</td>
<td>1.33 (.49)</td>
</tr>
<tr>
<td>4</td>
<td>2.66 (.75)</td>
<td>3.03# (.58)</td>
<td>3.09 (.60)</td>
<td>2.50 (.94)</td>
<td>2.50 (.92)</td>
<td>1.75* (.45)</td>
</tr>
</tbody>
</table>

Wilcoxon signed-rank test comparing consecutive years’ means: ** p < 0.01; * p < 0.05; #p = 0.05

Because we knew that changing the interview protocol might change the nature of the data beginning in Year 2, transcript coders noted when participants’ responses were coded on a particular dimension or sub-dimension before the new prompt was asked (if it was asked at all). To test whether the significant differences in RT and PS in Table 3 were attributable solely to the change in interview protocol, I repeated the comparison of scores in years 1 and 2 on only those interviews in which RT or PS were coded before their respective prompts. Results of the Wilcoxon signed-rank tests indicated that the differences remained significant (though for RT, it was at the p < .05 rather than the p < .01 level).
These two sets of results are intended to illustrate the kinds of findings that the VHQMI coding analysis yielded. In the final section, I consider these and other possible applications of the instrument in terms of validity and feasibility.

Discussion

In this paper, I have introduced and described the development of an interview-based instrument for assessing the sophistication of individuals’ visions of high-quality mathematics instruction. This work was motivated by the following research question: What trajectories do mathematics teachers’ and others’ visions of high-quality mathematics instruction follow over time in settings in which leaders are promoting models of instruction aligned with mathematics education research? Within each of the dimensions for which I developed a rubric, participants’ descriptions were rarely as elaborated as those in the literature, and often differed at least in terms of underlying rationale—or, as previously described, the functions underlying the forms (Saxe et al., 1999; Spillane, 2000). These differences are exactly what the VHQMI rubrics are intended to capture.

The results of the initial coding analyses described in this paper suggest that the VHQMI can be applied with sufficient reliability, and that the rubrics do represent trajectories of development, since those results reflect the kind of growth that the participating district reform efforts and the MIST project were designed to promote.

Validity

Space limitations prevent the inclusion of a thorough validity assessment. A more detailed report of the properties of the VHQMI instrument, including the relationships of VHQMI scores to other measures of knowledge, perspectives, and practice, has been reported elsewhere (Munter & Correnti, 2014). Here, I will limit my argument to three points. First, given
that most of the VHQMI dimensions and sub-dimensions draw from previous research in mathematics classrooms, to the extent that the rubrics’ foci and leveling aligns with one’s own instructional vision, the VHQMI instrument likely meets standards for content validity. Second, completing trajectories to model the development of instructional visions by drawing on data collected from school and district personnel whose professional learning was being pursued and supported supplements the vision of practice described in the research literature with those of practitioners, such as the examples of individuals’ trajectories highlighted above. Additionally, the contributions of experienced project members (e.g., the task force described above) increase the rubrics’ face validity.

Third, as noted in the previous section, average VHQMI scores seem to conform to both the chronology and history of district reform efforts. Specifically, as shown in Table 3, across the four years, VHQMI scores improved, suggesting some potential impact of professional development and other reform efforts on individuals’ instructional visions. And, as illustrated in Figure 9, on average, scores appeared to differ by district in predictable ways, given what we knew about teachers’ and others’ prior experience with reform-oriented curriculum materials and professional development.

That differences in VHQMI scores and patterns of change in those scores mirror approaches to professional development and other characteristics of the four districts suggests that the VHQMI instrument can be used to assess and track relevant aspects of district-wide improvement efforts. In the next subsection I discuss the feasibility and benefits of doing so.

**Feasibility and Benefits**

Methodologically, the means by which teachers’ and others’ knowledge, perspectives, or beliefs can be documented and characterized are numerous. Choosing to employ interviews to
assess (change in) visions of high quality mathematics instruction undoubtedly requires more time and expertise on the parts of both interviewers and coders than would using pencil-and-paper assessments or electronic surveys. But the nature of much of what the VHQMI instrument is intended to assess—as previously characterized by a form-function distinction (Saxe et al., 1999; Spillane, 2000)—renders other means of data collection inadequate. On a survey, participants might be asked to explain “why” they named a certain aspect of the classroom as something they would look for. But only in an interview can data collectors—with a participant still there and available to respond to further probing—determine whether they have thoroughly surfaced a participant’s perception of the function underlying an identified form. If such distinctions matter to the researcher, then interviewing is perhaps (at least currently) the only effective approach.

Of course, assessing individuals’ instructional visions based on what they say in an interview setting does not necessarily help one distinguish instances in which participants “really mean it” from those in which they are repeating what they have learned to say but do not value. But, because the VHQMI instrument employs a set of semi-structured interview prompts, what it is intended to assess is inherently about talk. That is to say that, although the VHQMI instrument is used to characterize participants’ current perceptions of high-quality practice, it is, more directly, an assessment of the ways that they can currently articulate those perceptions when given the opportunity. If responses are interpreted as such, then the extent to which respondents value the aspects of practice they describe is less relevant.

After all, as noted previously, new forms of practice may follow new ways of talking about practice. And, from the standpoint of typical professional development initiatives in school districts, it is through talk that the bulk of teacher learning is expected to happen. In most MIST
districts, for example, a significant component of improvement efforts required mathematics coaches to articulate and promote a particular instructional vision through regular, school-based collaboration and conversation among colleagues. In such cases, the VQHMI instrument may be used to assess what coaches might be able to articulate for school-based groups of teachers. Similarly, it may be helpful to assess what district leaders might be capable of describing in professional development sessions or what principals might be able to articulate after teaching evaluations.

**Application of the VHQMI Rubrics**

Just as Simon (1995, p. 142) argued that “anticipating students’ learning processes” enables teachers to more effectively support students’ mathematics learning, having models of trajectories of the development of sophisticated instructional visions may aid those who seek to foster such learning on the part of school personnel.

However, it is important to stress two aspects of the models. First, as explained above, they are built with a particular agenda for mathematics learning in mind, one that values learning mathematics with understanding (Carpenter & Lehrer, 1999), including the development of multiple strands of mathematical proficiency (Kilpatrick, Swafford, & Findell, 2001) and forms of mathematical practice (Common Core State Standards Initiative, 2010). To achieve such goals for learning, this agenda privileges instruction that affords students opportunities to engage in problem-solving, sense-making, building on current understanding, connecting mathematical ideas, authoring their own learning, communicating, and reflecting (Carpenter & Lehrer, 1999; Gravemeijer, 1994; 2004; NCTM, 2000; Putnam, Lampert, & Peterson, 1990; Sfard, 2003). Whether such instruction is termed “complex” (Cohen & Lotan, 1997; Boaler & Staples, 2008), “ambitious” (Lampert, Beasley, Ghousseni, Kazemi, & Franke, 2010), “responsible” (Ball,
2010), or simply “high-quality,” what is certain is that, in terms of what school district leaders choose to promote (and support teachers in accomplishing), such a vision of instruction is atypical (Elmore, 2004). Therefore, the trajectories I have proposed may be most applicable in settings in which a similar agenda for mathematics learning as that described above is espoused by those designing for and supporting teachers’ and others’ learning—though this need not be at the scale of the MIST project.

Second, the VHQMI rubrics do not represent “hypothetical learning trajectories” for teachers and others in the sense that Simon (1995) described such trajectories for students, since they do not include guidance on how those individuals’ learning should be supported through professional development (Clements & Sarama, 2004). A primary goal of the MIST project was to produce an empirically grounded, generalizable theory of action for improving the quality of mathematics teaching at scale (Cobb & Smith, 2008). Although such a theory of action would not specify instructional sequences in which to engage teachers, principals, and coaches, it would provide concrete guidance to school and district leaders concerning the structures and resources that are necessary for supporting teachers in developing new forms of practice (Cobb & Jackson, 2011). When combined with such a theory of action, then, the learning goals for school and district personnel represented by the top levels of the VHQMI rubrics and the progressions those rubrics model may form a district-level version of a hypothetical learning trajectory that assists school and district leaders in anticipating and supporting teachers’ and others’ learning processes.

With the above caveats in mind, assessing individuals’ visions of high-quality mathematics instruction with the instrument described in this paper is potentially useful for a number of purposes. In the larger MIST study, VHQMI analyses contributed to our annual
feedback to district leaders. In our end-of-year reports on how districts’ improvement efforts were actually playing out and our suggestions for improving the design in the following year, we used the data generated by employing the VHQMI instrument as a measure of district-wide progress in areas of interest. Whether a district’s professional development focus was on talk in the mathematics classroom or the use of complex mathematical tasks, we could note whether we had observed any changes, across role groups, in how individuals described such aspects of the classroom in their instructional visions. Additionally, members of the project have worked to understand the extent to which changes in VHQMI scores relate to changes in other aspects of teachers’ (and others’) knowledge and practice, including assessments of instructional quality, mathematics knowledge for teaching (Hill and Ball, 2004), and other interview- or survey-based assessments (e.g., Munter & Correnti, 2014; Wilhelm, this issue; Wilhelm, Munter, & Jackson, 2014).

But these are likely only a fraction of the potential uses of the instrument. VHQMI assessments could also serve as measures of individuals’ levels of expertise in examinations of the emergence and impact of professional social networks; as a characterization of the level of discourse in which a group of collaborating teachers engage; as a pre-post assessment of “uptake” from professional development experiences; as a progress monitoring tool in teacher training programs; or as an intermediate variable in the identification of mechanisms by which district reform efforts do or do not impact the nature of classroom mathematics instruction. To make use of the VHQMI instrument in these and other analyses pertaining to teachers’ and others’ learning, the prompts and rubrics could be adapted (e.g., elaborating additional sub-dimensions, using only a subset of the rubrics, posing the prompts in electronic survey form,
etc.) to suit both research questions and the contexts in which researchers or teacher educators are working.

**Equity**

Before concluding, I wish to acknowledge one last limitation of the instrument. As stated previously, my intent was to be able to adequately characterize teachers’ and others’ VHQMI with as few dimensions as possible. Although I conducted my analyses essentially through the lens of literature-based constructs, I limited the instrument’s dimensions to those that most frequently captured MIST participants’ ways of describing instruction (i.e., role of the teacher, classroom discourse, and mathematical tasks, and student engagement). Consequently, articulations of a number of arguably crucial dimensions of mathematics classroom learning and teaching cannot be directly assessed with the instrument—perhaps most importantly, in light of my own, more recently developing perspective, equity.

To be clear, it is my perspective that instruction cannot be “high-quality” if it is not equitable, including affording all students access and opportunity to engage in and contribute to meaningful mathematical activity in spaces in which they feel they belong—where their racial, gender, and other identities are affirmed and strengthened. While some equity-promoting practices are represented in the VHQMI rubrics (e.g., shared authority, emphasis on conceptually oriented talk, tasks with multiple entry points), others are not (e.g., “culturally ambitious” practices aimed at critical consciousness; Waddell, 2014). In the future, it may be beneficial to incorporate (a multi-faceted definition of) equity as an additional dimension for characterizing instructional visions.

**Conclusion**
Recently, Hiebert (2013) argued that “improving classroom teaching is one of the most serious and urgent problems facing mathematics educators” (p. 46). Approaching this problem as one of supporting teachers’ (both pre- and in-service) and others’ learning (as Hiebert suggested) requires the explication of not only the end goal (i.e., high-quality mathematics instruction as defined in the literature), but also the progressions that individuals might take to get there. In the absence of such models it is difficult to approach the work of studying or supporting teachers’ learning developmentally. The VHQMI instrument described in this article provides one such “roadmap” in that it identifies potential milestones in individuals’ perceptions and articulations of practice as their instructional visions and enactments evolve over time.
References


doi:10.1080/00461520.1996.9653265


Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class
discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education, 44*(4), 646–682. doi#:10.5951/jresematheduc.44.4.0646


Michaels, S., O’Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and


doi:10.1207/S1532690XC11802_01


doi:10.1080/07370000701301125


doi:10.1080/10986060802229675


doi:10.3102/00028312033002455


Appendix

VHQMI Rubrics
### Developing Visions of Instruction

**Level:** 4) Teacher as more knowledgeable other

**Potential ways of characterizing teacher's role:**
- **Influencing classroom discourse:** Suggests that the teacher should purposefully intervene in classroom discussions to elicit & scaffold students' ideas, create a shared context, and maintain continuity over time (Staples, 2007).

| Examples |
|------------------|------------------|
| Teacher plays a proactive role in supporting/ scaffolding students' talk: “When [teachers] pose a question and a student answers, they don’t say yes this is how it is always done. They ask the kids to explain how they came up with the answer, ask for other students to explain how they came up with the answer, present all the ideas to the student and ask them if these are good procedures for answering types of problems like this and talk about student preference; ‘Do you like one way more than another and does this way make sense?’ — so that the kids can build their own frame of reference to the material.” |

**Attribution of mathematical authority:** Suggests that the teacher should support students in sharing in authority (Lampert, 1990), problematizing content (Hiebert et al., 1996), working toward a shared goal (Hiebert et al., 1997), and ensuring that the responsibility for determining the validity of ideas resides with the classroom community (Simon, 1994).

| Examples |
|------------------|------------------|
| Teacher uses students’ explanations, responses, questions, and problems as lesson content (Fraivillig et al., 1999); “Students should be involved in the learning process as far as asking questions and being able to maybe actually give examples and working them and talking to the teacher about them.” |

**Conception of typical activity structure:** Promotes a “launch-explore-summarize” lesson (Lappan et al., 1998), in which a) the teacher poses a problem and ensures that all students understand the context and expectations (Jackson et al., 2013), b) students develop strategies and solutions (typically in collaboration with each other), and c) through reflection and sharing, the teacher and students work together to explicate the mathematical concepts underlying the lesson’s problem (Stigler & Hiebert, 1999).

| Examples |
|------------------|------------------|
| Teacher facilitates the process of students working together to solve problems and then share explanations. “Students are the ones doing most of the work and they’re defending their answers and they are challenging each other.” |

---

**Level:** 3) Teacher as facilitator

**Potential ways of characterizing teacher's role:**
- **Influencing classroom discourse:** Describes the teacher facilitating student-to-student talk, but primarily in terms of students taking turns sharing their solutions; hesitates to “tell” too much for fear of interrupting the “discovery” process (Lobato et al., 2005).

| Examples |
|------------------|------------------|
| Teacher asks students about other students’ work, or to be prepared to ask their own questions about other students’ work (Hufferd-Ackles, Fuson, & Sherin, 2004) but does not articulate a rationale for such teacher moves in terms of supporting the development of a discourse community; instead, the teacher “facilitates” question/answer time after student presentations, without intervening to highlight key mathematical ideas and possibly concluding by telling the class the “correct” solution/strategy. |

**Attribution of mathematical authority:** Supports a “no-tell policy”: Stresses that students should figure things out for themselves and play a role in “teaching.” Suggests that if students are pursuing an unfruitful path of inquiry or an inaccurate line of reasoning, the teacher should pose a question to help them find their mistake, but the reason for doing so focuses more on not telling than helping students develop mathematical authority. Is open to students developing their own mathematical problems, but these inquiries are not candidates for paths of classroom mathematical investigation.

| Examples |
|------------------|------------------|
| “[Students are] not waiting all the time for the teacher [to] come and spoon-feed them but doing investigating on their own, coming up with ah-ha!s on their own or coming up with ‘what if this?’—that’s when I think they’re really learning.” |

**Conception of typical activity structure:** Promotes a “launch-explore-summarize” lesson (Lappan et al., 1998), in which a) the teacher poses a problem and possibly completes the first step or two with the class or demonstrates how to solve similar problems, b) students work (likely in groups) to complete the task(s), and c) students take turns sharing their solutions and strategies and/or the teacher clarifies the primary mathematical concept of the day (i.e., how they “should have” solve the task).

| Examples |
|------------------|------------------|
| “It depends on the topic that you are presenting. Some topics the teacher may have to present and then other topics it’s better to let the kids explore.” |

---
### 2) Teacher as monitor

**Influencing classroom discourse:** Suggests the teacher should promote student-to-student discussion in group work.

- Teacher should encourage students to “ask each other for help”
- “If a kid understands the way to explain it better than I do, I give the floor to them as long as it makes sense.” A student who “gets it” should come to board and teach—“Having a kid who's really good at the math, but who's still at their [peers'] level, sometimes they can explain it a little bit better [than the teacher]”
- “If the students are going down a direction that looks like it's a dead-end, “the teacher needs to circle the wagons, regroup, 'Oh guys this is not working out. We need to back up cause, cause we're going the wrong way. So let's back up. Let's try this a different way.’”

**Attribution of mathematical authority:** Suggests a view of teacher as an “adjudicator of correctness” (Hiebert et al., 1997). Students may participate in “teaching” but only as mediators of the teacher's instruction, adding clarification, etc. If students are pursuing an unfruitful path of inquiry or an inaccurate line of reasoning, the teacher stops them and sets them on a “better” path.

- The teacher should “show [students] examples [of] how to do it and why are they doing it, what is the purpose of it. Then, do the facilitation, walk around, see the group work”; “Students acting as leaders”; “Students should be teaching each other”
- Teacher should be “up, moving around, looking at students' work, helping them as they work”; Teacher should not “lecture the whole time”

**Conception of typical activity structure:** Promotes a two phase, “acquisition and application” lesson (Stigler & Hiebert, 1999), in which a) the teacher demonstrates or leads a discussion on how to solve a type of problem, and then b) students are expected to work together (or “teach each other”) to use what has just been demonstrated to solve similar problems, while the teacher circulates throughout the classroom, providing assistance when needed.

- “Teacher should be mathematically correct”; “no steps skipped”
- “Explain why & how it's used in everyday life, not just formulas”

### 1) Teacher as deliverer of knowledge

**Influencing classroom discourse:** Focuses exclusively on teacher-to-student discourse. Considers quality of teacher's explanations in terms of clarity and mathematical correctness.

- “Teacher should be mathematically correct”; “no steps skipped”
- “Explain why & how it's used in everyday life, not just formulas”

**Attribution of mathematical authority:** Suggests that the responsibility for determining the validity of ideas resides with the teacher or is ascribed to the textbook (Simon, 1994). (This includes insistence that teachers be mathematically knowledgeable and correct.)

- Teacher “should explain it so it makes sense”; “The teacher needs to be factually accurate”; “I would look and see that the teacher seems to know what he or she's talking about.”
- Teacher “should answer all student questions”; “If there's a misconception is the teacher correcting it or letting it go?”

**Conception of typical activity structure:** Promotes efficiently structured lessons (in terms of coverage) in which the teacher directly teaches how to solve problems. Periods might include time for practice while teacher checks students’ work and answers questions, but this is likely quiet & individually-based with no opportunity for whole-class discussion. Description suggests no qualms with exclusive lecture format.

- Teacher provides clear instructions, clear assignment, examples shown, students being walked through a problem”; “Most of the time the teacher teaches it and the students take it in. If they have any question about it, then they should feel free to ask.”
- Teacher has a task to accomplish—to present the lesson planned—and must see that it is accomplished without digressions from, or inefficient changes, in the plan (Thompson, 1984).

### 0) Teacher as motivator

- Looks to see “whether the teacher has the dynamics that the kids need.” Looks for “The teacher's enthusiasm”; “Some teachers are obviously more charismatic than others.”; “It is more about being an entertainer than it is a teacher.”
- Looks for teacher “making connections to students. Some people are just naturally very good at teaching.”

---

*Figure 11. VHQMI Rubric: Role of the Teacher (abbreviated)*
<table>
<thead>
<tr>
<th>Level</th>
<th>Patterns/structure of Classroom Talk</th>
<th>Nature of Classroom Talk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Description</strong></td>
<td><strong>Example(s)</strong></td>
</tr>
<tr>
<td>4</td>
<td>Promotes whole-class conversations, including student-to-student talk that is student-initiated, not dependent on the teacher (Hufferd-Ackles, Fuson, &amp; Sherin, 2004); promotes developing &amp; supporting a &quot;mathematical discourse community&quot; (Lampert, 1990).</td>
<td>Suggests that classroom talk should be conceptually oriented—including articulating/refining conjectures and arguments for explaining mathematical phenomena—for the purpose of supporting students in “doing mathematics” and/or spawning new investigations.</td>
</tr>
<tr>
<td></td>
<td>• &quot;A child will come up and show their work and a different child explain what they did to solve it. And then the children actually question each other, ‘Well I didn’t see where you got that three from, can you show me where did it come from?’ or, ‘Why did you do it this way instead of this way?’ They're communicating with each other and the teacher's more of a facilitator.”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Promotes whole-class conversations (about ideas, not just whole-class lecture or task set-up), but description places the teacher at the center of talk, likely doing most of the prompting and pressing, or calling upon students/groups to take turns presenting their strategies.</td>
<td>Insists that the content of classroom talk be about mathematics (e.g., asking questions, providing explanations), but description of such talk either (a) characterizes talk that is of a calculational orientation; or (b) fails to specify expectations for the nature/quality of the questions, explanations, etc.</td>
</tr>
<tr>
<td></td>
<td>• Describes a view of students asking questions of one another's work on the board, but likely at the prompting of the teacher, where students usually give information when probed by the teacher with some volunteering of thoughts (Hufferd-Ackles, Fuson, &amp; Sherin, 2004).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• &quot;[In] classroom discussion I would expect the teacher to throw out a question as a facilitator and then I would expect the students to somewhat lead that discussion. This is how we got to this; this is your next step. This is the next step.' in these different groups raising their hands and telling what the next steps and how to solve a problem are. So I would think that again the teacher would be the facilitator and the students would be kind of leading the discussion.”</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Promotes student-student discourse but describes it exclusively in the context of small group/partner work (if there’s mention of whole-class discussion, it’s characterized only as an option, not a vital element)</td>
<td>Insists that the content of students’ classroom talk (with each other) be about mathematics, but provides no description of content (i.e., does not specify things such as questions and explanations).</td>
</tr>
<tr>
<td></td>
<td>• “The larger the discussion, usually the harder it is for them to discuss. More than likely it's teacher led discussion... I like for them to actually discuss with their partner”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Students should “engage with each other”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• “Students should ask each other questions instead of asking the teacher”</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Describes traditional lecturing and/or IRE (Mehan, 1979), or IRF (Sinclair &amp; Coulthard, 1975) dialogue patterns. (Note that this can occur in a “whole-class” setting, but is not considered a genuine whole-class discussion.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• “Most of the time the teacher teaches it and the students kind of take it in... there’s not a lot of room for debate on the math because, you know, this is it!”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Students “could answer the questions you ask” (i.e., in an IRE pattern, the teacher’s questions are clear and answerable).</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 12. VHQMI Rubric: Classroom Discourse (abbreviated)*
<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Example(s</th>
</tr>
</thead>
</table>
| 4     | Emphasizes tasks that have the potential to engage students in “doing mathematics” (Stein, Grover, & Henningsen, 1996; Smith & Stein, 1998), allowing for “insights into the structure of mathematics” & “strategies or methods for solving problems” (Hiebert et al, 1997). | • Students should be engaged in challenging questions that have ambiguous or multiple routes to a solution in order to generate multiple solution paths and strategies for discussing/comparing, thus promoting students' flexibility in applying problem-solving strategies (Russell, 2000)  
• “Questions that pertain to their lives around them or connected to things they've done in previous days… and require the kids to learn a concept not just by being told what it is and how to do it but to actually think about what it is they were doing and then coming up with the why or ‘Oh look, this worked for all these problems, so is this gonna work for all of our problems?’… do some critical thinking” |
| 3     | Emphasizes tasks that have the potential to engage students in complex thinking, including tasks that that allow multiple solution paths or provide opportunities for students to create meaning for mathematical concepts, procedures, and/or relationships. “Application” is characterized in terms of problem solving. However, tasks described lack complexity, do not press for generalizations, do not emphasize making connections between strategies or representations, or require little explanation (Boston, 2012). Instead, they emphasize connections to “the real world” or “prior knowledge.” Reasons for multiple strategies are not tied to rich discussion or making connections between ideas. | • Tasks should have “more than one solution or maybe different ways to approach it so that different ideas are accepted and could be possible”  
• “A problem can be solved different ways, because there are different ways of thinking and kids need to know that there's not just one set way to do things”  
• “Have multiple entry points for students, multiple solution tasks that require children to really think and put a lot of information together in order to answer the question”  
• “Open-ended so it doesn't have a right answer, and it talks about how things fit together instead of what the answer is (e.g., ' me some starting and ending points that could be a vector of positive 5')”  
• “I would look for tasks that accessed some sort of prior knowledge yet took the kids a little bit further to build on that knowledge.”  
• “I want to see a lot of different ways of doing the same thing… some kids can get past the visual and they're into the abstract mode much quicker and then they don't want to waste their time and be bored.” |
| 2     | Promotes “reform”-oriented aspects of tasks without specifying the nature of tasks beyond broad characterizations (e.g., “hands-on,” “real world connections,” “higher order”), and without elaborating on their function in terms of providing opportunities for “doing mathematics” (Stein, Grover, & Henningsen, 1996; Smith & Stein, 1998). “Application” is characterized in terms of “real world” context and/or students being active. | • “Hands-on activities, instead of doing worksheets… maybe build something or work it out with some kind of a model… the application of what they've learned is really important.”  
• “higher order thinking problems with application”  
• “bring in the outside world to try to get the kids engaged”  
• “not doing straight book work” (instead, “cutting out puzzle pieces and making two puzzles to prove Pythagorean's Theorem”)  
• “Tasks should include “time to move and use those manipulatives and things.” |
| 1     | Emphasizes tasks that provide students with opportunity to practice a procedure before then applying it conceptually to a problem (Hiebert et al, 1997)                                                                 | • “First is to understand what the concept is, and what the formula is, and how to do it in terms of the numerical way. Second is applying it… if it's put into a word problem” |
| 0     | Either (a) does not view tasks as inherently higher- or lower-quality; or (b) Does not view tasks as a manipulable feature of classroom instruction                                                                 | (a) Depends on the teacher, “whatever works for them”; “Depends on the class”; “The thing that actually gets them to start asking questions”  
(b) “We're supposed to be using the CMP book which is pretty much, this is what you do and here's what the teacher should say and it even tells you how it should run.” |

*Figure 13. VHQMI Rubric: Mathematical Tasks*
<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Example(s)</th>
</tr>
</thead>
</table>
| 2     | Specifies WHAT students should be doing, using typical reform language and focusing primarily on the organization/structure of the activity, without describing the nature of classroom activity in content-specific ways (i.e., describes a 'non-traditional' classroom, full of activity, but does not specify how the activity is specific to mathematics). If reasons WHY particular forms of activity are important are provided they are not in terms of supporting students' participation in doing mathematics. | • “The majority of the kids engaged in thinking and doing investigation, switching from traditional teaching to inquiry based”  
• “Students should be up moving around using manipulatives”  
• “Students should be presenting”  
• Students should be doing "investigations, experiments…making their own graphs and comparing them with people at their tables, where they’re actually doing stuff together. They’re not just taking notes” |
| 1     | Stresses the importance of students being engaged and "on-task", either taking for granted the quality of classroom activity (i.e., students should be doing whatever the teacher asked), or specifying traditional classroom activities as what should take place. Response either (a) stresses THAT students should be engaged and participating in classroom activities (i.e., on-task, paying attention), without specifying WHAT those activities should be, or (b) describes nature of classroom activity as traditional classroom activity. | (a) “student engagement… if students are working and actively participating”; “I think that being able to see that everybody is on task and hearing the questions that the teacher asks and listening to the students' responses, I can tell from there just from that if they really understand what's going on”  
(b) “student engagement… they’re participating and writing down what they need to be, taking notes, listening or appear to be listening. I mean not sleeping, not getting up and walking around the room. They're in their seats, they’re working as a group, you know or maybe they’re working in groups, maybe they're not, but you know they seem to be focused.” |

*Figure 14. VHQM Rubric: Student Engagement in Classroom Activity*