Proceedings

Conference on Mathematical Modeling

Teachers College
Columbia University

October 14, 2013

Editors:
Benjamin Dickman
Andrew Sanfratello

Table of Contents

What Counts in Mathematics (and Other) Classrooms? A Framework for Looking at What Matters, and Thoughts About How One Might Use These Ideas for Professional Development

Alan H. Schoenfeld
University of California

This article describes the evolution of an analytical scheme that describes the dimensions of powerful mathematics classrooms—classrooms that produce students who are powerful mathematical thinkers. Such classrooms engage students with coherent and connected mathematics; they provide *all* students opportunities to do mathematical thinking. They help students develop a sense of mathematical agency and authority, and they employ formative assessment, finding out what students understand and building on those understandings. I discuss how this scheme can be used as the basis for professional development, and how a school district might take up a full-fledged professional development agenda.

Keywords: dimensions of productive teaching, professional development, cognitive demand, equity, agency, mathematical identity, formative assessment

I'm really delighted to be here today. It's always good to come back to Teachers College, but today is special: we're celebrating twenty-five years of Henry Pollak's tenure at TC. Note that this is only Henry's second career. I can't imagine what he'll do in his third career over the next twenty-five years—but I'm sure it'll be spectacular. Thank you, Henry, for providing an excuse for me to come back to TC again.

Let me begin by giving an overview of what I'm going to discuss this evening. What I want to think through with you is what counts in mathematics classrooms, and how we can use an understanding of what counts in professional development. I love mathematics; I think it's a wonderful subject for its beauty, its power, and the fun one can have doing it. I spend my time trying to figure out how to make a world in which students experience mathematics in the rich and powerful ways that I was lucky enough to experience it. My hope is that we can ultimately move the system so that students get to experience mathematics as the kind of sense-making activity that it can and should be. That's been the overarching issue for me career-wise.

For the past half dozen years or so, I've been thinking about the apparently simple question of what makes for a productive mathematics classroom, and how you can capture that in an analytic scheme. In what follows I'm going to frame the big question, then trace the painful evolution of our analytic scheme, and then talk about how, once one has some sense of what the issues are, we might use those understandings as a basis for professional development.

It's evening and a lot of people have had a long day, so rather than bore you all the time, I'll also show you some classroom videos. I did ask Bruce Vogeli if we could have some popcorn for the movies, but it doesn't seem to have materialized. Do remember that there's beer right outside the door. At the end, if you're still sober—or maybe not—we'll talk. First, a tip of the hat to my sponsors: many thanks to the National Science Foundation (NSF) and the Bill and Melinda Gates Foundation for the funding that enables me to do this work.

Let me tell you what the big framing is, and then narrow down on tonight's substance. I should explain that even though I was trained as a pure mathematician—which means I'm theoretical and supposedly don't deal with the real world—there is an applied side to me. In particular, when I make statements, I like to have data to back them up.

That's not necessarily the case in the political arena. For example, give or take a couple years the math wars were fought between roughly 1990 and 2000. Although the flaws in the traditional curriculum were well known and there had been some small scale studies suggesting that curricula along the lines of the NSF-supported curricula could be valuable, the fact is that there was no hard evidence either for or against the Standards-based curricula between 1990 and 2000. The NCTM standards were written in 1989, and the request for proposal for the standards-

base curricula came out soon afterwards. Alpha versions of the curricula were developed and taught, roughly between 1990 and 1995; beta versions between 1995 and 2000. There was very little research on, or evaluation of, those curricula through 2000. The first book on results, edited by Sharon Senk and Denise Thompson, came out in 2003. In sum, there was little research-based evidence, pro or con, between 1990 and 2003. Yet, the math wars flourished—an absence of data didn't stop the arguments.

Clear evidence began to emerge in 2003, at which point we could say: "Hey, look, it's now well documented. If you teach a balanced diet of skills, concepts, and problem solving, your students will do as well on a test of skills as students who got nothing but a diet of skills—and they'll beat the pants off of them when it comes to concepts and problem solving."

In short, I like having evidence for the things I believe in. So how does that relate to tonight? The thing is, we all think we know what's important in teaching. The odds are that my friends and I would all give the same thumbs-up or thumbs-down after watching an hour of instruction. But, do we have any evidence that the things we like make a difference? Not really.

I want to be in a position to say, "Here are the things we value—and, we have empirical evidence that in classrooms that score high along these dimensions, the students who emerge are powerful mathematical thinkers. They do well on real tests of thinking and problem solving." That's the game I want to play. So I got together with Bob Floden at Michigan State and my friends from the Shell Center at Nottingham to address these issues. Rather than do all of mathematics, Bob suggested we focus on algebra at first—it's a more manageable topic. In general, we'd like to know: what makes for powerful mathematical thinking? To get a manageable handhold on the topic, we'd look at word problems in algebra—but we'd abstract that eventually. Of course, we need to trace the impact of the teaching, so we need some good tests of mathematical thinking and problem solving. Happily, we have the MARS/Balanced Assessment tests, and they're not bad. So what we needed to do is to build a frame for capturing what goes on in mathematics classrooms along the dimensions that count. Then we can proceed empirically—we can look at two or three zillion classrooms and see if the classrooms that score high along these dimensions are producing students who are powerful mathematical thinkers, and those that score low on the analytic scheme are producing students who score low on tests of thinking and problem solving like the MARS tests. The relevant scheme is in Figure 1.

That's the issue at hand. So the question is: how do we describe the classroom practices we think produce powerful mathematical thinkers? Can we create a scheme that captures every kind of classroom you can imagine in a comprehensible way, so that then we can look for the relationships between scores on this scheme and scores on good math tests?

I want to start with some images of practice, to give us concrete referents for our abstract discussions. The first is the TIMSS geometry video. A second is of Cathy Humphreys doing the border problem, some of which I showed earlier today. These are both widely available.

I know a lot of you have seen the TIMSS video, so I'll describe the most important aspects of it very briefly. The key pedagogy it demonstrates is what are called *IRE sequences*. The teacher asks the question, or **I**nitiates. Typically, the student has a short amount of time to **R**espond; then the teacher **E**valuates what the student said.

[The TIMSS 8th grade geometry video was played. Here is a brief summary]. In a typical excerpt, the teacher points to a pair of intersecting lines with the four angles marked A, B, C, and D; D is marked as 70 degrees (Figure 2).
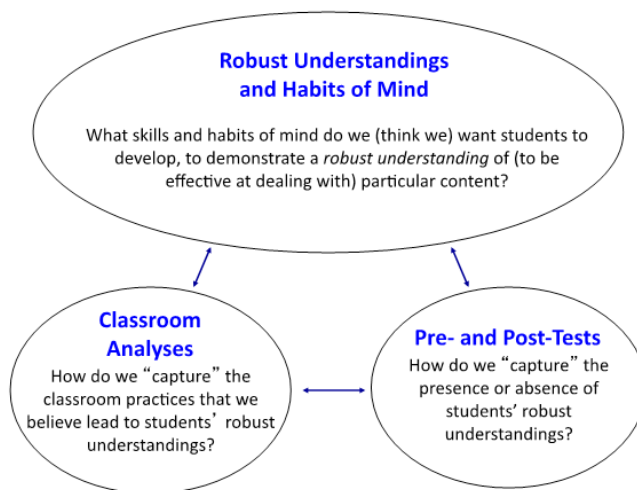

Figure 1.

The teacher then asks a rapid-fire series of questions and reinforces the answers:

Figure 2.

"Angle *A* is vertical to which angle? (students volunteer the answer "*D*") Therefore, Angle *A* must be? . . . (students volunteer the answer) . . . Seventy degrees. Go from there."

"Now you have supplementary angles. What angle is supplementary to . . . Angle *A*?

B is, and so is? . . . *C*. Supplementary angles add up to what number? (Students volunteer the answer) One hundred and eighty degrees. So if you know one is seventy, the other one has to be a hundred and ten. Go from there."

The rest of the video is similar. The teacher frames a question with a short answer and calls on a student who has a short amount of time to respond. If the response is correct, then the teacher repeats it and moves on to the next short answer question; if it is not, then he calls on someone else or provides the right answer. [The video was stopped here.]

That's about as much we need to see. You notice that if a student took more than three milliseconds, the teacher passed right on to the next one. There's no opportunity for the students to think at all.

Earlier today, I showed a bit of a movie that I want to show more of now. The teacher in this tape is Cathy Humphreys, a master teacher in California. (The tape is available as part of Jo Boaler and Cathy Humphreys' 2005 book *Connecting Mathematical Ideas*.)

[The video was played. Here is a brief summary.] Humphreys asks the students to figure out, mentally, how many squares are on the border of a 10-by-10 grid. After she asks, hands go up silently. When almost all of the students have indicated that they have an answer, she asks the students to discuss their answers at their tables; then she re-convenes the class. She asks if the students want to shout out the answer (36), which they do. Then she asks if anyone had first gotten 40 as an answer. Some hands go up, and a student explains how she had first computed 4×10, but then realized that some of the border squares overlapped. Humphreys continues by asking if anyone had gotten 38. Yes, said some students, and one explains how. It is clear from the tape that there is no opprobrium attached to arriving at a wrong answer; the idea is to be alert to the possibility of mistakes, and fix them.

Humphreys then asks a series of students to explain their reasoning. One by one they go to the front of the class, explaining their ways of visualizing the sum. As they do, she writes down each method, next to the student's name (A's method, B's method, etc.). After each student has explained his or her method, Humphreys checks in with the class, to make sure that they understand what the presenter has said. Then, with the 10-by-10 grid thoroughly discussed, Humphreys asks the students to imagine a 6-by-6 grid, and to use A's method, B's method, and so on, to compute the number of squares in the border of that grid.

It's obvious where this will be going: the various methods will be applied to find the number of border tiles in an *N*-by-*N* grid. This will yield a number of different algebraic expressions: $[4(N) - 4]$, $[2(N) + 2(N - 2)]$, $[N + 2(N - 1) + (N - 2)]$, $[N2 - (N - 2)2]$. And, shouldn't they all be equal? [The video was stopped here.]

Having seen these clips, let's return to the task I set out for myself. If I'm to be able to characterize classrooms in general, I have to be able to come up with a scheme that covers everything from the TIMSS tape to the Cathy Humphreys tape, and everything else you've ever seen in a classroom. The challenge is, can you come up with a scheme that's comprehensive *and* comprehensible? That's the goal. The idea is to have a small number of dimensions that explain all the variance. In any talk, one can only give a broad outline of the ideas. For more detail see Schoenfeld (2013).

For those of you who know my problem solving work, the claim that I made back in 1985 was that there are four categories of behavior you have to look at in order to understand success or failure in problem solving. You've got to look at the solver's (i) knowledge base, (ii) use or non-use of problem solving strategies, (iii) metacognition (monitoring, and self-regulation), and (iv) belief systems. The claim was that if you look at any problem solving episode, the cause of success or failure will reside in one or more of those four categories—and that you're not going to need any more categories than that.
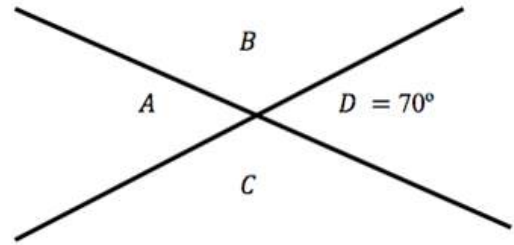
The number four is important because of the magic number seven plus or minus two. In some classic research, George Miller (1956) discovered that 7 ± 2 is the maximum number of things that people can hold in short term memory. If I give you a scheme with twenty-eight things, you need a long checklist to use it; but if I mention four or five key ideas, you can keep them in mind. Thus my goal for looking at classrooms was to come up with a scheme whose dimensions were important, relatively independent, and relatively small in number. Let me tell you about the twists and turns on the way to discovering the scheme.

The last time I was here I talked about the theory of decision making that I was working on at the time—the question being: what drives how and why people make the choices they make while they're in the middle of complex activities like teaching? The book that resulted from that work, *How We Think,* came out about two years ago. It identified what's important in teachers' decision making. So it stands to reason, then, that if you understand teachers' decision making, you know exactly what to look for in classrooms, right?

Wrong. Using the word "model" in the strict computational sense, I can model an hour of teaching down to every utterance that the teacher makes. The only problem is it may take me three years. If I'm looking for a useful scheme to capture what's going on in classrooms—and I literally want to look at hundreds, if not thousands, of classrooms—I can't afford the luxury of modeling in detail. We need a scheme that's going to do all the things I mentioned before *and* preferably be operable within, say, twice real time. I'd like to be able to go into an hour-long class, take notes, turn those notes into a set of scores, and use those scores, all within two hours or so. The theoretical approach that I had was great for modeling, but it was far too complex for "twice real time" activities.

That didn't stop me from trying. My poor graduate students at the time were in great pain as we tried to code all those video tapes. But finally, I realized that it just doesn't work; we had to start over again. It was time to return to the literature to see what we could find. There's a bunch of stuff. There's Charlotte Danielson's (2011) "framework for teaching"; there's CLASS (CLassroom Assessment Scoring System) by Bob Pianta (Pianta, La Paro, & Hamre, 2008); there's PLATO (Protocol for Language Arts Teaching Observations), which is a language arts system but looks at what's going on in classrooms (Institute for Research on Policy Education and Practice, 2011). There's MQI, developed by Heather Hill (Hill, Charalambous, & Kraft, 2012), and the UTeach protocol (Marder & Walkington, 2012). There's more. Why invent something new, when these are all available?

The answer was that none of these had all the attributes we wanted. Some, for example, are really, really good at very specific things. If you're interested in classroom discourse, there are schemes that address it—but I wanted something comprehensive and schemes that focused on discourse didn't meet that criterion. The same is the case for inventories of teachers' mathematical content knowledge, and so on. At the other end of the spectrum, there are comprehensive schemes. Charlotte Danielson's scheme covers everything, but I dare you to tell me what the half dozen key dimensions are. Recall that I wanted something that was distilled into a small number of powerful dimensions.

In short, the extant schemes helped us understand what to look for, but didn't help us to focus and distill. So, we were back to building our own. Here I want to show the simple outline of what we tried to do at our first pass [shows slide].

The slide shows three main dimensions of mathematical activity: Access (who gets to participate), Accountability (what's expected of the students), and Disposition (what's likely to emerge from the ways students engage with the mathematics). We code each of them in four categories: with regard to the quality of the mathematics, the learning of mathematics, the classroom community, and the individual learner. This may or may not be exactly right, but it certainly seems manageable. Except it's not, because what I've just shown you is the outline. When we got down to detail [shows slide with the coding detail] we had dozens and dozens of codes, and the scheme was really difficult to use. So, we put this approach on the shelf, and, again, looked for something else.

The idea was to see if we could break the logjam of having too much to look at—so, what about events of particular interest or importance? What happens when your ears perk up, because something of consequence happened? Could we capture the importance of those "events of interest?" We tried to categorize those, and came up with three focal aspects of classroom activity. First, there's the general classroom environment. That has to do with managerial issues—whether the students are unruly or sitting there waiting to be called on but petrified, or active participants, and so on. You can talk about those things independent of whether you're in a math class, a social studies class or anything else. Second, there are the classroom mathematical norms. How do people play the game of mathematics? Is it about filling in the blanks? Is it about sense making? Is it about something in between? Third,

because we were interested in algebra, we wanted to see the specifics of doing algebra. So we built rubrics for those things. Here [showing slides] was the rubric for general classroom stuff, here was the rubric for general mathematics, and here was the rubric for the specifics of doing algebra. Yes, you're shaking your heads appropriately. Once again, we got lost in detail—this stuff isn't easy.

So we tried yet again! I won't tell you about that version—you can thank me later. We went back to the drawing board, armed with the insights from the various approaches. Well, there was the notion that there's a natural flow of classroom episodes. That's been part of my research from the very beginning. So why don't we keep that? And then why don't we say what were the important things within those episodes, what were the things that really mattered? And then say, during those times, what was important?

The idea was as follows. You sit in the classroom. You take structured field notes. Then you chunk your notes into episodes. They capture the bare bones structure of what happened; for example: the teacher started with classroom business, then discussed homework, then lectured on the main substance, then had the students practice what they'd been shown. There were rules for parsing the lesson into episodes that were pretty straightforward. Next, for each of those episodes, you have to say what was important. (This was the descendent of the "events of interest" from a previous scheme.) They were called facets—and in one episode, you might have three or four facets (representing types of interesting things) and then you would code each of these facets.

Well, you can guess the punch line: this was more structured, but it, too, had so many things to look for that it was too unwieldy to be usable. Happily, though, we had reached the point where we were focusing on the right stuff. That is, all the things we were sure were important were there to be looked at; the problem was that they were arrayed in a giant mosaic in which the key dimensions, whatever they were, were buried. So I did a classically mathematical thing, which was to sort them into equivalence classes. For each item, I asked, "What is this an instance of?" And, suppose we cluster together all the things that are of the same type. Can we identify that cluster as an important dimension of a lesson?

The answer was yes, and I'm about to lay out for you the dimensions of the **T**eaching for **R**obust **U**nderstanding of **Math**ematics (TRU Math) scheme for classroom analysis. But first, a technical note. For the purposes outlined at the beginning of this talk, showing that classrooms with certain properties produced students who were powerful mathematical thinkers, we would have to assign scores along each dimension. And, because classrooms have different activity structures at different times—whole class discussions, small group work, individual seat work, student presentations—we would have to have separate scoring rubrics for each of those activity structures. That's just technical detail, however. In what follows I'm going to focus on the five main dimensions that emerged from our analyses. I'll introduce them via five questions, then elaborate on each.

Key Questions for Math Classes

1. What were the big ideas, and how did they get developed?
2. Did students engage in "productive struggle," or was the math dumbed down to the point where they did not?
3. Who had the opportunity to engage? A select few, or everyone?
4. Who had a voice? Did students get to say things, develop ownership?
5. Did instruction find out what students know, and build on it?

Dimension 1 is concerned with the mathematics at the heart of the lesson. What was the quality of the mathematics? What were the big ideas and how did they get developed? Were the students engaged in mathematical sense-making? The importance of this dimension should be obvious: the quality of the mathematics being discussed is a key determinant of what the students walk away from the class with.

Dimension 2 concerns what is known in the literature as "cognitive demand." We think of it as the opportunity for students to engage in productive struggle. The key question is: how much intellectual work are the students actually getting to do? The reason this is important is that there's a large body of research that says that when students find things difficult and ask for help, the modal response from teachers is to say "Here, do it this way"—thereby removing the challenge for the student. This is a problem: if the teacher is the one doing the heavy lifting, then the students aren't developing their mathematical muscles. The real art of teaching is in providing the appropriate scaffolding for students. You don't want to spoon-feed them, and you don't want them at sea. If a student is confused, can you

provide just enough structure or guidance so that the student can now grapple productively with the mathematics in front of him or her?

Dimension 3, broadly speaking, is the classroom equity dimension. We all know teachers who, because they want their lessons to proceed smoothly, consistently call on the three sharpest students. That way they get the answers that allow them to speed the lesson along. But, the question is, what about the remaining $(n-3)$ students? A really equitable classroom provides opportunities for every single student to engage with the mathematics at his or her level productively.

Dimension 4 can be thought of as the "discourse" dimension. Technically, it addresses what we call agency, authority, and identity: do students come to see themselves as people who can do and explain mathematics? Think about the difference between the two videos I showed you. In the first tape, students got to say one word that completed the teacher's sentence. In the second tape, the students got to go up to the front of the class and explain their thinking. They answered questions. The teacher also labeled each of the methods—Student A's method, Student B's method, etc. As a result of this practice, students were developing the ability to explain mathematics and receiving the recognition for having done so. Ultimately, that becomes part of one's mathematical identity. Phil Daro, who's one of the authors of the Common Core State Standards and on the advisory board for our project, said that if he had one measure he could use to predict the power of the mathematics classroom, it would be the number of times during the hour that a student had the opportunity to say a second sentence in a row of explanation. That's a pretty powerful measure.

Dimension 5 is the formative assessment dimension, I talked about that a lot this morning, so I'll just say here that it's important: really effective lessons meet students where they are, building up what's right, and addressing what isn't.

As I mentioned above, these are the core dimensions of the TRU Math scheme. By the time this article appears in print, the most recent version of the TRU Math scheme (including the rubrics for different classroom activity structures) will be posted on the Algebra Teaching Study (http://ats.berkeley.edu/) and the Mathematics Assessment Project (http://map.mathshell.org/materials/index.php) web sites, along with a guide to understanding and using the scheme.

Although we don't have zillions of analyses yet (we have dozens) it seems that with enough practice on the part of the coder, the scheme is usable. You can sit in the classroom, take notes, and code your notes in somewhere between two and three times real time. We're in the process of gathering the data that I hope will allow us to address the question: do classrooms that work powerfully along these dimensions produce students who are powerful mathematical thinkers? My bet is yes, but as I said, I believe in data.

Also, we think that, although we built TRU Math for use in algebra classrooms (the full scheme contains a dimension focusing on the kinds of classroom practices that should enable algebra students to make sense of, and use algebraic reasoning to solve, complex algebraic word problems) it should be straightforward to build analogous content-specific modules for geometry or other mathematical content areas. So we think TRU Math will work for all mathematics classes. I'm willing to bet that the scheme will be general, in that classrooms that do well on dimensions 1 through 5 (and attend to whatever the relevant content specifics might be) will produce students who are powerful mathematical thinkers. So now we come to the bottom line question: If you think those five dimensions are important, what are you going to do about it?

My goal is to help teachers develop the capacity to engage more productively in each of these five dimensions—*all* of these five dimensions. My best guess is that you need for all of these things to be in place, for instruction to shine. So how do you go about this? Well, you don't do what I've been doing this evening, namely, talk about a scheme for evaluating classrooms. The TRU Math scheme is a research tool, and of necessity we have to evaluate what goes on in the classrooms we're studying. But when you talk about working productively with teachers, anything that comes across as evaluation turns people off. That's not what we're about. The bottom line of the research is that the five dimensions in the scheme (rich mathematics, cognitive demand, equitable access, agency and identity, formative assessment) represent important aspects of productive mathematics classrooms. If you take that as a given, then you can switch gears and say: How can we turn those ideas into a productive tool for conversations with teachers? Happily for me, some of my current graduate students are experienced and talented professional developers; they've done the bulk of the detailed work, including collaboration with local school districts.

The issue is: Can you come up with a set of questions for conversations with teachers that generate rich and productive conversations about planning a lesson, reviewing the lesson, and thinking about what to do next? So you take the core questions and modify them, to spark good conversations.

- Question 1, related to mathematical rich mathematics (focus and coherence): How did the mathematical ideas in this unit (or this course) get developed in this lesson (or this lesson sequence)?
- Question 2, related to productive struggle: What opportunities do students have to make their own sense of mathematical ideas?
- Question 3, related to equitable access and participation within the classroom: Who does or doesn't participate in the mathematical work that the class does, and how do they participate? And then, of course, how can they be supported in doing more?
- Question 4, related to agency and identity: What opportunities do students have to explain their own ideas and respond to each other's mathematical ideas?
- Question 5, related to formative assessment: What did we learn about the students' current mathematical thinking and how can we build on it?

Now each of those is in essence a topic sentence. It's the introduction to a conversation. To keep the conversation going, you can expand each question into a family of questions. For example, question 1, "How do important ideas get developed?" can be expanded to: "What are the goals for the lesson? What connections exist between this lesson and ideas in the past or what you're building for? How do we take the procedures that are used in the lesson and make sure they're justified and connected with important ideas? How do we see and hear students engage with those ideas in the class?" And so on. For cognitive demand, you can ask, "What opportunities exist for the students to struggle with mathematical ideas? How can we support that engagement? How have I been responding when the students get stuck? Are there ways I can avoid answer-giving but scaffold student work so that the students still have plenty to work on? Similarly, how do you think about access? What might we do to get all of the students to the point where they're engaging in mathematics productively? How do you think about supporting their mathematical agency? Who generates the mathematical ideas that get discussed, and who evaluates them? How deeply do students get to explain their ideas? And how does the teacher respond to those ideas? Is it by questioning, by proving, by soliciting responses from other students?" All of these questions are ways of opening conversations between teachers, between teacher and coach, in individual teacher reflection.

Now, I'll be the first to say that it's really hard to build the skills and understandings I've been talking about here. I was at a meeting a few weeks ago with representatives of school districts serving more than a million kids. The person who preceded me ended her overview of the issues by saying, "By this time next year, we will be fully Common Core compliant." Can you guess what my first word was? "Bull." We have to recognize the reality that the kinds of things we're talking about are extremely challenging, and they will take time, even under the best of circumstances. Moreover, we have to think systemically. All of you know the administrator who goes into a mathematics classroom where kids are arguing about a problem in small groups, and says, "A good math classroom is a quiet classroom, with the teacher in charge. This noise is disruptive." End of story, end of progress. The issues we're talking about have to be tackled on a school-wide, district-wide basis, because incoherent or mixed messages negate progress.

So, what do you do? Here's the grand plan. If Oakland Unified School District, which is next door to Berkeley and was one of the first districts taken apart under No Child Left Behind, had the full resources, this is what we'd be doing with them. Earlier today I discussed the formative assessment lessons (FALs) that the Mathematics Assessment Project has developed—see (http://map.mathshell.org/materials/index.php). These are designed to be embedded about two-thirds of the way through a content unit, to provide rich feedback about what the students have made sense of and what they need to work on. Say your school year has nine units. You pick nine FALs that match the content being covered, and embed a relevant FAL two-thirds of the way through each unit.

Here is our plan; we are hoping to get the funding that would enable us to implement it. The first thing we're going to do is get together all of the administrators in the district: the superintendent, deputy superintendent, district support staff, and all building principals and vice principals for instruction. And we're going to teach them a formative assessment lesson or two. They're going to be the students in those lessons, so they get to experience what it's like to be in a classroom where the focus is mathematical understanding, and a mechanism for attaining that understanding is rich discussion. After this experience we give the administrators a tool they can use for their

classroom observations. Phil Daro created a tool cleverly called "The Five by Eight Card"—you can guess how big it is. It contains a bunch of questions for classroom observers to address when they visit classrooms. It's focused on classroom discourse. "Are the students getting to explain? Are the students listening to each other's explanations?" This shifts the administrator's focus away from "Do I like how the teacher is running the class?" and turns it to "Are the students talking about the mathematics in meaningful ways?"

That's support for administrators. Here are our plans for teachers. Imagine that through the course of the year, at the beginning of each unit, we teach the relevant FAL to all of the teachers who are going to be teaching that unit. After teaching the FAL (with the teachers as students), we debrief. We'll use the TRU Math framework as the over-arching structure for debriefing. We'll go through the dimensions: What were the big ideas? How did they fit into what you're doing? Did you guys get to struggle in an appropriate way? Did everybody here have an opportunity to engage? And we'll do that using the extended set of questions we discussed earlier. My hope is at the end of nine iterations over the course of the year, the teachers will have internalized those questions.

In addition, imagine planning or debriefing sessions between teacher and coach, between fellow teachers planning or reviewing the formative assessment lesson. The frame for these conversations is the question version of TRU Math.

Of course, you're wondering at this point, "How would teachers have the time for such conversations?" Let me tell you about what we made work in Berkeley some years ago. We had been doing professional development (PD) with Berkeley teachers on a voluntary basis. At the end of the first year, the teachers said, "We like this but we don't like the fact that it's after school and off the clock. It's a voluntary effort, but it should be part of our work week. We're going to go to the teachers union and ask them to negotiate with the district. We'd like to lengthen the school day on Monday, Tuesday, Thursday, and Friday. That way we can have a short Wednesday, allowing for professional development in the afternoons, while maintaining the same number of contact hours with students. Finishing classes earlier on Wednesday will allow us to do PD as part of our workday. That will give us the time to collaborate with each other."

So now imagine that when the teachers are going to teach the formative assessment lessons, they have planning time with their colleagues. Then they teach the FALs as lesson study lessons, where other teachers are free to come in and watch. The lessons are videotaped, and the teachers now have Wednesday afternoon PD time to debrief. They do so using the TRU Math questions, of course: "Did the math come out the way we thought it might? Did the students have the opportunity for productive struggle?" And so on. In this way, TRU Math becomes a collaborative mechanism by which teachers debrief on their teaching practices. Do that for five years and you can really change teaching.

How do you start on a five year plan, since five dimensions are too much to handle at once? In year one you can focus on two main things. First, some of the math in the Common Core is new—so one of the things you want to address, as soon as possible and every year until it becomes comfortable, is the core content. "Let's talk about the math" is a natural entry point—and the FALs provide a good way to address much of the key content. Second, there's discourse: "Let's focus on opportunities for your students to talk mathematics." We may spend the whole first year just focusing on those two things. Why? Well, they're hard, for one thing. For another: once you get discourse going, you have opportunities to see what kinds of struggle kids are engaging in and who's doing it; and, when your students are talking math, you have greater opportunities to hear what they understand and what they don't. That is, an early focus on dimensions 1 and 4 (the math, the discourse) lays the groundwork for deeper attention to dimensions 2, 3, and 5 (productive struggle, equity of access, and formative assessment). I could say more, but I've already gone over my time. So, let's open things up for discussion. Any question is fair game.

Audience member: I have a question about differentiation, both for students and for professional development for teachers. In a classroom that has such a great equity, there are some students who really get it; what opportunities will they have to go further—either wider or more vertically? And the same thing for teachers, I guess, and the same thing for supporting those at the bottom. So what about differentiation?

Schoenfeld: I want to talk about that in two ways. The first is to discuss a particular form of equity-oriented instruction known as Complex Instruction. The concept was developed by Elizabeth Cohen and Rachel Lotan at Stanford (Cohen & Lotan, 1997). A recent book describing how it works is *Strength in Numbers: Collaborative Learning in Secondary Mathematics* by Ilana Horn (2012). I don't have the time to describe the whole process here, but a core aspect is that a lot of work is done in small groups, where all of the group members are responsible for (a) sorting out the mathematics, and (b) making sure that everyone in the group truly gets it. If a group says it's

ready for the teacher to look at their work, the teacher selects one student randomly, and that student is responsible for explaining the group's work. If the student doesn't succeed, the teacher—with no negatives attached—says she'll return, and the group goes back to work, making sure that the selected student "gets" the math.

Complex Instruction (CI) would not be very effective if the math was drill-and-kill procedural mathematics. Part of what makes it powerful is the notion of "group-worthy" problems. Such problems demand thought—but they also have multiple entry points, so students can approach them in different ways. Thus they support rich conversations and the sharing of ideas. When CI is done effectively, every student can make a meaningful contribution to the group's mathematical work. You don't have to have different problems for different students in order for all the students to be engaged in ways that are mathematically meaningful and appropriate.

On that score, let me tell you about a different technique that I learned from Phil Daro. In some ways, what Phil does is the analog of what you can see in the Japanese TIMSS algebra tape. There, a class works on one problem for a whole hour, and the teacher works through a range of ways to work through it. Here's what Phil does.

He starts with what would normally be a five-minute algebra problem:

Train A leaves the station at fifty miles an hour. Three hours later, Train B leaves on a parallel track going sixty miles an hour. When does Train B overtake Train A?

In a traditional algebra class, we would devote just a few minutes to this problem. If Train B has traveled t hours (the amount of time it takes to overtake Train A) then it has traveled $60t$ miles while Train A has traveled $50(t + 3)$ hours. Solve the equation $60t = 50(t + 3)$ and you're done. Well, you may or may not be done. Doing the algebra gets the answer, but it leaves a lot unexamined.

Phil breaks the class into pairs, giving them the problem and telling the students they can work it in any way they prefer. He does think-pair-share, after which the students make posters showing their solutions. What you'll find in an average algebra class is that half the students—even though this is an algebra class—wind up making tables. For Train A they enter 50, 100, 150, 200, . . . For Train B they may start with a 60 in the third box, maybe the fourth box, 60, 120, 180, 240, . . . They continue until the two numbers match—which is at either 15 or 18 hours, depending on when you start counting time. A significant proportion of the students do graphs, so you get the two graphs that have positive slope and that, depending on how you set the graphs up, either intersect at $t = 15$ or $t = 18$. Some students do the algebra. Some, the clever ones, go, "I never use algebra. The first train had a 150 mile head start. The second one was catching up at a rate of ten miles an hour, so it took 15 hours to catch the first one."

What Phil does is to start with the least sophisticated solution, which in this case is the table. He picks one of the posters with a table on it, and says, "Hey, guys, why don't you tell us about what you did with your solution?" The class talks about the table and he raises a whole bunch of issues, so they really work through the table carefully. How did you know where to start the second number? Did you have any sense of where it might catch up? Did you see any interesting patterns? And so on. He works the table to death. Then he turns to a poster with the graph on it. He has the students explain the graph, and then says, "I want to take every question we asked about the table and ask it about the graph. You said it was catching up by ten miles an hour because you saw that in the differences. Where do you see the 'catching up' in the graph?" Many of the students won't have made a connection like that before, but once you know to look for it, the pattern of decreasing distances in the graph becomes visible. "Some of you have the graph of Train A starting at $x + 0$, and Train B at $x = 3$; some of you have the graph of Train A starting at $x = -3$, , and Train B at $x = 0$. What's negative time ($x = -3$) in this context? And how does that relate to starting time on the tables we looked at? Some of you have a point of intersection at $x - 15$ and some at $x = 18$. How do those values relate to the values we found on the tables? Then he goes to the algebraic solution, and works through it in similarly. "How do you know how far each train has gone after t hours? Is that t hours after Train A or Train B left the station? Where is that in the table, where is that in the graph?" Then: "We saw catching up in the table and in the graph. Where do you see the catching up in the algebra?" Finally, he works through the non-algebraic solution, connecting it to the other three.

So, Daro has taken a five-minute problem and blown it up into a day-and-a-half of classroom discussion. We all know teachers' response to that move: the curriculum doesn't allow the time for me to do that! Is what Phil did impractical, if not impossible?

Actually, no. If you look at the content that he covered, he did a week and a half's worth of content—tables, graphs, algebra, and connections across them—in two days. In working through the problem, he reviewed everything the students need to know about the topic. But equally important, this way of approaching the math meets issues

of differentiation head on. The students who have what we consider the most mathematically primitive solutions get to see how their solutions are the foundations for more sophisticated solutions, and how things begin to fit together. The students who had zoomed straight to the algebraic solution got to deal with some questions they wouldn't have thought about, and they got to see how their approaches to the problem were connected mathematically to the other approaches. It's a win for everyone. The answer to the problem of differentiation is create a mathematical context that is so rich and connected that *everybody* has something to learn there.

Audience member: I really appreciate the five dimensions you mentioned, and the fact that you said it was going to be hard and take time. I get the impression you're talking about in-service teachers. So I was wondering if you have given any thought to the learning trajectories of beginning teachers and how it fits in with these dimensions.

Schoenfeld: How many hours have you got? There's a theory of teachers' developmental trajectories—see Chapter 8 of my book *How We Think.* To be honest, we have a big problem in this country because we don't treat teachers as professionals—either in the pre-service phase or when they're out in the field. One of the first things I did after moving to Berkeley was to work on revamping our teacher preparation program. We had a good program, but one in which few research faculty were involved. I volunteered to teach a problem solving course as part of it, and did for a few years. One day I ran across a student who'd been out for four years or so. I asked her if the problem solving course had been useful. "Not for the first two years," she said; "I was just trying to survive. But, once I felt comfortable in the classroom, I started drawing on the things we did in that class. They're now a core part of what I do as a teacher."

I thought that said a lot. We have to help our beginning teachers survive, but also give them tools for growth—things to think about on an ongoing basis, so that when they get their "sea legs" they have fundamentals to draw upon. So, we built the program as a two-year program. The students take the same core courses as our Ph.D. students (including my problem solving course, which I'm teaching this semester). They have four different placements over their four semesters, to experience significant diversity in the groups they teach. They belong to faculty research groups, and a significant part of their work, including their Master's papers, is to reflect on what they've done. That builds an experimental, reflective habit of mind, which is what will keep them fresh as they grow through their careers.

Audience member: Okay, I have a question that may sound strange. Throughout your career you've emphasized metacognition. I don't have to convince you how important it is. But I don't see it in your scheme. Where is it?

Schoenfeld: It's in two places. For the teachers, it's built into the question form. The questions are designed to be asked in planning, and then in reflecting on how the lesson went. For the students, it resides largely in dimensions 1 and 4. Dimension 1 is formally the "mathematics dimension," which asks, is the mathematics treated in a way consistent with our best interpretation of the Common Core State Standards? The practices are all about building productive habits of mind—and to me, reflecting on your work is central in that arena. Dimension 4 is the "discourse and identity" dimension. It focuses on opportunities to speak mathematics. I didn't present the details here, but high scores on this dimension mean that students have, and take advantage of, the opportunity to propose ideas, explain them, have them critiqued and refined—that the discourse community subjects ideas to rigorous and thoughtful evaluation. It's in such interactions that monitoring and self-regulation become productive habits of mind.

I could say more, but it's nine o'clock and you've all been very patient. Thanks so much for the opportunity to come, and for the conversation.

References

Boaler, J. & Humphreys, C. (2005). *Connecting mathematical ideas: Middle school video cases to support teaching and learning.* Portsmouth, NH: Heinemann.

Cohen, E.G. & Lotan, R.A. (Eds.). (1997). *Working for equity in heterogeneous classrooms: Sociological theory in practice.* New York: Teachers College Press.

Danielson, C. (2011) The Framework for Teaching Evaluation Instrument, 2011 Edition. Downloaded April 1, 2012, from http://www.danielsongroup.org/article.aspx?page=FfTEvaluationInstrument.

Hill, H., Charalambous, C., & Kraft, M. (2012). When inter-rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher 41*(2), 56–64.

Horn, I. (2012). Strength in Numbers: Collaborative Learning in Secondary Mathematics. Reston, VA: NCTM

Institute for Research on Policy Education and Practice (2011). PLATO (Protocol for Language Arts Teaching Observations). Stanford, CA: Institute for Research on Policy Education and Practice.

Marder, M., & Walkington, C. (2012) UTeach Teacher Observation Protocol. Downloaded April 1, 2012, from https://wikis.utexas.edu/pages/viewpageattachments.action?pageId=6884866&sortBy=date&highlight=UTOP_Physics_2009.doc&

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review, 63*, 81–97.

Pianta, R., La Paro, K., & Hamre, B. K. (2008). *Classroom Assessment Scoring System.* Baltimore: Paul H. Brookes.

Schoenfeld, A. H. (1985) *Mathematical problem solving.* Orlando, FL: Academic Press.

Schoenfeld, A. H. (2011). *How We Think: A theory of goal-oriented decision making and its educational applications.* New York, NY: Routledge.

Schoenfeld, A. H. (2013) Classroom observations in theory and practice. *ZDM, the International Journal of Mathematics Education, 45:6–7–621.* DOI 10.1007/s11858-012-0483-1.