Summative and Formative Assessments in Mathematics Supporting the Goals of the Common Core Standards

Alan H. Schoenfeld

University of California, Berkeley

Alan H. Schoenfeld is the Elizabeth and Edward Conner Professor of Education and affiliated professor of Mathematics at the University of California, Berkeley. His email address is alans@berkeley.edu.

Draft, November 18, 2012.

Running head: Summative and Formative Assessments

This manuscript is part of a proposal for a special issue of *Theory into Practice* titled "Assessment of Complex Thinking in Mathematics, Reading, Science and History" edited by Kadriye Ercikan.

Abstract

Proficiency in mathematics involves a great deal more than mastery of facts and procedures. It involves having rich domain knowledge including: access to problem solving strategies, having a productive disposition and domain-specific beliefs, being a strategic thinker, and being appropriately metacognitive. This broad set of mathematics goals is central to the Common Core State Standards for Mathematics.

In the context of high-stakes examinations, teachers focus their attention on what is tested. Hence it is critically important for the assessments produced by the two national consortia, the Smarter Balanced Assessment Consortium (SBAC) and the Partnership for Assessment of Readiness for College and Careers (PARCC), to truly reflect the values of the Common Core Standards – and for teachers to be provided with classroom assistance in meeting those standards. Test specifications and sample assessment items from the two consortia are discussed, as are the prospects that the assessments will be positive levers for change.

For more than 20 years the Mathematics Assessment Project has focused on the development of summative assessments that emphasize the mathematical processes and practices discussed above. More recently the project has focused on the creation of “formative assessment lessons” or FALs to help teachers build up student understandings through the use of formative assessment. This paper describes our recent work. It illustrates the scaffolding our lessons provide, to support teachers in posing rich tasks and being prepared to respond to the range of student conceptions (and misconceptions) that their students are likely to produce.

Introduction

The United States stands at the threshold of significant changes in mathematics assessment, both in terms of what kinds of understandings are assessed and in terms of the increasing homogeneity of mathematics assessments, nationwide. These changes reflect the continued evolution of the "standards movement," which can be dated back to the of the National Council of Teachers of Mathematics' (NCTM) production of the *Curriculum and Evaluation Standards for School Mathematics* in 1989 and to a radical change in the national high stakes accountability context due to the "No Child Left Behind" legislation passed by Congress in 2001. They have the potential to be truly consequential for mathematics education in the US. As explained below, within a few years the vast majority of American students will be taking one of two high stakes examinations, both of which are intended to represent the mathematical values represented in the Common Core State Standards for Mathematics, or CCSSM (Common Core State Standards Initiative, 2010). To the degree that the assessments represent the values in CCSSM, and to the degree that high stakes assessment drives instruction, mathematics teaching in the US will be much more focused and coherent than it has been over the past quarter century.

This introduction briefly describes the dual evolution of mathematics standards and the national testing context. With that context established, we then examine some typical current test items, and some of the items that represent the assessments being produced by the two national assessment consortia, the Smarter balanced Assessment Consortium (SBAC) and the Partnership for Assessment of Readiness for College and Careers (PARCC). As will be seen, issues of alignment with the CCSSM remain; but, assuming that these can be worked out, the new assessments portend significant change. That being the case, the question is how to prepare students and teachers for such change. One way of doing so is being explored by the

Mathematics Assessment Project (MAP[1]), which is producing 100 “Formative Assessment Lessons” aimed at supporting classroom practices consistent with the CCSSM. A sample lesson is described, and some of the practical challenges faced by mathematics educators are discussed.

The evolution of standards, 1975-2010

The “standards movement” began when the National Council of Teachers of Mathematics, in reaction to national curricular chaos and the perception that the U.S. was falling significantly behind the other industrialized nations in the preparation of Science, Technology, Engineering, and Mathematics (STEM) majors (see, e.g., the National Commission on Excellence in Education’s 1983 report, *A Nation at Risk*), produced a volume of mathematics standards representing NCTM’s vision of high quality mathematics curricula and evaluations. The 1989 *Standards* were revolutionary, in that they were the first national curriculum document that gave significant attention to mathematical *processes* (what one does while engaging in mathematics) as well as to *content* (e.g., algebra, geometry, probability, statistics, and data analysis, etc.). The *Standards* declared that at each grade band (K-4, 5-8, and 9-12), significant attention must be given to mathematics as problem solving, to mathematics as communication, to mathematics as reasoning, and to mathematical connections. This approach was reified in NCTM’s 2000 volume *Principles and Standards for School Mathematics*, which reiterated the emphasis on mathematical processes in the original *Standards,* and added a focus on the use of mathematical representations.

This “process-oriented” view of mathematics was grounded in decades of research on mathematical thinking and problem solving, which emphasized not only the content of

[1] The MAP project gratefully acknowledges funding from the Bill and Melinda Gates Foundation.

mathematics, but the productive ways in which people went about learning and using their mathematical knowledge. For example, Schoenfeld (1985, 1992) characterized someone who is mathematically *proficient* (which is presumably the goal of mathematical instruction) as someone who, in addition to having a substantial amount of mathematical knowledge, also has access to productive problem solving strategies, has learned to be effective in metacognitive terms (using his or her resources efficiently), and who has developed productive beliefs and dispositions about mathematics and her- or himself as a doer of mathematics. Focusing on the elementary grades (but with implications at all grade levels), the National Research Council's 2001 volume *Adding it Up* identified five intertwined strands of mathematical proficiency – see Figure 1.

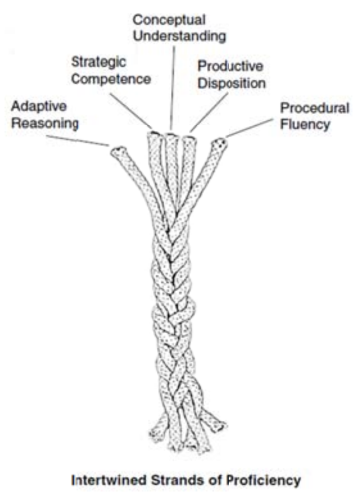


Figure 1. (National Research Council, 2001, p. 5)

The *Common Core State Standards* represent the natural evolution of the work discussed above. Given the context of their creation (see below), there are some differences; but overall, one sees more detailed specification of content and the continued emphasis on process, now reframed as mathematical *practices*). Perhaps the most important change is that in the content descriptions in the CCSSM, content is specified for each grade through grade 8, rather than in broad grade bands. This means that content progressions are much more fine grained than previously – e.g., the specifics of rational numbers and fractions are laid out in a way that leads to discussions of proportionality and linear functions. (The authors of CCSSM stress that focus and coherence are main virtues of the Standards.) This will have significant entailments for curricula and assessment, in that both are far more determined than previously. (For example, "seventh grade content" will be taught and assessed in grade seven, whereas it might have been possible before the CCSSM for that content to have been taught or assessed in grade six or eight as part of the middle school "grade band.") Where previous documents referred to "processes," the CCSSM refer to "Standards for Mathematical Practice." However, the authors make their heritage clear:

> "The Standards for Mathematical Practice . . . rest on important "processes and proficiencies" with longstanding importance in mathematics education. The first of these are the NCTM process standards of problem solving, reasoning and proof, communication, representation, and connections. The second are the strands of mathematical proficiency specified in the National Research Council's report *Adding It Up*: adaptive reasoning, strategic competence, conceptual understanding (comprehension of mathematical concepts, operations and relations), procedural fluency (skill in carrying out procedures flexibly, accurately, efficiently and appropriately), and productive disposition (habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficacy)."

The eight mathematical practices highlighted in the CCSSM are that students will:

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments…

4. Model with mathematics
5. Use appropriate tools strategically
6. Attend to Precision
7. Look for and make use of structure
8. Look for and express regularity in repeated reasoning.

The challenge for assessment has been, and will continue to be: Is it possible to reliably capture ability at such practices?

The curriculum and assessment context, 1975 - present

In the 1970s and through the 1980s, a small number of states had state-wide mathematics standards; a smaller number (e.g., California, New York, and Texas) had assessments that were aligned to those standards. This began to change with the issuance of the NCTM Standards in 1989; across the nation, various states adopted mathematics standards. In curricular terms, there was significant entropy. On the one hand, it has been said that there are 15,000 independent school districts in the United States – in the sense that, in the 1970s and 1980s, school districts had tremendous latitude in selecting educational goals and the means to achieve them[2]. Top-down constraints, if they existed, were only state-wide: formally, any two adjacent states could in principle follow their own independent rules (if they had any). On the other hand, there were real-world constraints. No matter where a school district was located, it had a limited number of textbook options, given that there were a relatively small number of major publishers producing textbook series. Moreover, the curriculum was somewhat constrained from the top down, in that

[2] There were some constraints. California is a "textbook adoption" state: districts could use whatever texts they choose, but the districts were only reimbursed by the State for the costs of those texts if the texts were on a state-approved list. And, of course, extant exams (such as the Regents Exam in New York State) exerted significant pressure for homogeneity.

school districts were aware of college entry requirements. (At minimum, college-intending students had to be ready for calculus, which meant that they would typically take the standard Algebra I – Geometry – Algebra II – Trigonometry – Precalculus course sequence. But, there was a very uneven patchwork of standards and (if they existed in a state) statewide assessments to judge student proficiency. In the 1990s, for example, the primary method of assessment in Oklahoma was a multiple choice test focusing on basic skills, while Vermont employed a portfolio assessment of students' collected work on extended problems. In large measure because of the longstanding American tradition of states rights, there was little coherence to the system and little mandate for it.

The situation changed with the passage of the "No Child Left Behind" act, of 2001[3]. To qualify for federal funding under NCLB, as it is known, each of the states had to institutionalize standards for mathematical performance, and to assess students on a regular basis. These exams were "high stakes:" students' promotion, teachers' salaries (and jobs), administrators' salaries (and jobs), and the very existence of schools and districts (which could be dismantled if student test scores failed to meet the increasingly stringent scoring requirements over a period of years) depended on test scores. Not surprisingly, most schools focused heavily on teaching to the tests[4], which were of highly variable quality. Given that each state had its own standards and assessments, the result was nationally institutionalized incoherence.

[3] No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425-2094, 2002.
[4] Some years ago Hugh Burkhardt coined the phrase "What You Test Is What You Get (WYTIWYG)" to represent this reality. Space does not permit a discussion of WYTIWYG; but see Barnes, Clarke, & Stephens (2000) and Bell & Burkhardt (2001).

The context continues to evolve, with significant changes being catalyzed by the federal $4 billion Race to the Top (RTT) fund[5], which was part of the 2009 American Recovery and Reinvestment act announced by President Obama and Secretary of Education Duncan on July 24, 2009. The constraints of RTT were that *consortia* of states, not individual states, would apply for funding. This constraint led the Council of Chief State School Officers and the national Governors Association to sponsor the Common Core state Standards Initiative, which produced the CCSSM. To date, forty-five states and three territories have adopted the CCSSM – thus establishing what is, de facto, a national set of mathematics standards.

In addition, the Race to the Top Assessment Program[6] "provided funding to [two] consortia of States to develop assessments that are valid, support and inform instruction, provide accurate information about what students know and can do, and measure student achievement against standards designed to ensure that all students gain the knowledge and skills needed to succeed in college and the workplace." Those consortia, the Partnership for Assessment of Readiness for College and Careers, or PARCC[7], and the Smarter balanced Assessment Consortium, or SBAC[8], each have enrolled about half of the states that have agreed to align themselves with the Common Core State Standards. As a result, there will no longer be a patchwork of 50 state assessments. With the exception of the students in the five states and one territory that have not signed up for RTT, students across the country will be faced with one of two assessments, constructed either by PARCC or SBAC, and ostensibly aligned with the CCSSM. Given WYTIWYG, and the fact that CCSSM standards and assessments will be given

[5] See http://www2.ed.gov/programs/racetothetop/index.html
[6] See http://www2.ed.gov/programs/racetothetop-assessment/index.html
[7] See http://www.parcconline.org/about-parcc
[8] See http://www.smarterbalanced.org/

at each grade K-8, there will be a degree of homogeneity in curricula and in assessments that is unprecedented in American history.

The nature of mathematics assessments, past and possibly future

As noted above, mathematics assessments across the US have varied widely from state to state. Here I provide an example from the California Standards Tests (CSTs) as an example of what has been the reality in one state, and contrast this with a more rich assessment of proficiency in the same content area. I then discuss the item specifications and sample items from the two national assessment consortia.

Figure 2 contains a representative eighth grade algebra problem from the CST[9].

What is the y-intercept of the graph of $4\boldsymbol{x} + 2\boldsymbol{y} = 12$?

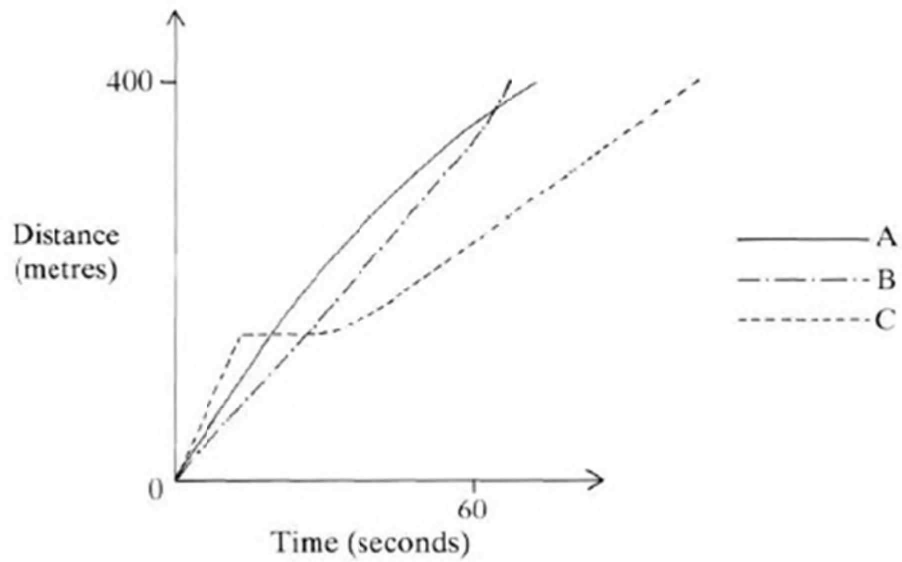
(A) - 4 (B) - 2 (C) 6 (D) 12

Figure 2. A released CST problem from the 8th grade algebra I test

This task, like most of those on the CST, focuses on content knowledge. There are at least three straightforward ways to get the answer: by substituting $\boldsymbol{x} = 0$ into the equation and solving the resulting equation, $2\boldsymbol{y} = 12$; by writing the equation in the slope-intercept form $\boldsymbol{y} = -2\boldsymbol{x} +6$; and by writing it in the two-intercept form $\boldsymbol{x}/3 + \boldsymbol{y}/6 = 1$. In each case, the procedure is mechanical and the answer straightforward to obtain. Although content knowledge is assessed, it is hard to argue that the standards for mathematical practice are assessed in any meaningful way.

In contrast, consider the "hurdles race" task given in Figure 3.

[9] This is problem 23 from the Algebra I released problems from the California Standards test, downloadable from the California Department of Education at <http://www.cde.ca.gov/search/searchresults.asp?cx=001779225245372747843:gpfwm5rhxiw&output=xml_no_dtd&filter=1&num=20&start=0&q=released%20items%20california%20standards%20test>. If is typical of the level of difficult of the exam. More sample questions can be accessed at <http://starsamplequestions.org/starRTQ/search.jsp>.



The rough sketch graph shown above describes what happens when 3 athletes A, B and C enter a 400 metres hurdles race.

Imagine that you are the race commentator. Describe what is happening as carefully as you can. You do not need to measure anything accurately.

Figure 3. Hurdles Race. Swan, M., and the Shell Centre Team (1985), p. 42.
Reprinted with permission.

This question calls for interpreting distance-time graphs in a real-world context, a central component of mathematical modeling. A complete response includes:

- Understanding that a runner whose graph appears “to the left” of another is *ahead* at that point, having taken less time to travel the same distance. (Thus B wins the race);
- Understanding what points of intersection signify in this context (that two runners have run the same distance at the same time, so they are tied at that point in the race)
- Interpreting the horizontal line segment (the runner is not progressing, so – in the context of a hurdles race – must have tripped on a hurdle and fallen), *and*
- Putting all of the above together in a coherent narrative.

Equally, important, responding appropriately to this question calls for demonstrating proficiency at (at least) the first four of the mathematical practices highlighted above. *If* tasks of this level of complexity will appear on the two consortiums' assessments, then there will be significant changes in what is assessed (and, by virtue of WYTIWYG, what is taught) across the nation.

From this author's perspective, there is significant promise that the two assessment consortia can move things in very productive directions – but, progress is hardly guaranteed. There are various places where things can go wrong: in the specifications for the exams; in ways the specs are realized in the exams themselves; and in the grading, to mention just three.

The Consortia's Exam Specifications.

Here I think there are grounds for significant optimism[10]. The fundamental change in the SBAC assessments is that they will report four scores, not just one. Until now, a student's score in most assessments was a single number. This allowed people (including the students assessed) to say how well the student did, on both an absolute and on a comparative scale (most high stakes exams provide data that yield percentile scores), but they provide no information about what the student did or did not do well. (For example, did the student do well on algebra but not geometry, or vice-versa? Did he or she earn most of his points on procedural questions, on those that asked for extended chains of reasoning, or on some of both?) In contrast, the SBAC (2012, p. 19) test specs call for reporting four scores for each student, corresponding to the following claims about what the assessments reveal:

> *Claim #1, Concepts & Procedures:* Students can explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency.

[10] Full disclosure: I was lead author for the SBAC mathematics content specifications.

Claim #2, Problem solving: Students can solve a range of complex well-posed problems in pure and applied mathematics, making productive use of knowledge and problem solving strategies.

Claim #3, Communicating reasoning: Students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others.

Claim #4, modeling and data analysis: Students can analyze complex, real-world scenarios and can construct and use mathematical models to interpret and solve problems.

In the most optimistic reading, this bodes extremely well for standards-based mathematics. It is quite clear that a test like the California Standards Test, with only multiple choice problems focusing on concepts and procedures, fails to assess claims 2, 3, and 4 in a meaningful way. Extended problem solving tasks, of complexity not unlike the "hurdles race" task given above, populate the SBAC specifications. If such tasks make their way into the actual assessments, they will (by virtue of WYTIWIG) drive classroom instruction in the direction of the CCSSM. But there are risks; see below.

The PARCC assessment promises tasks of three types: (1) Tasks assessing concepts, skills and procedures, (2) tasks assessing expressing mathematical reasoning, and (3) tasks assessing modeling/applications (PARCC 2012, p. 14). This is broadly consistent with the approach taken by SBAC and the CCSSM. It is not clear from the documents available on the PARCC website (http://www.parcconline.org/) what the format for reporting student scores will be, so I was unable to determine whether there will be separate scores for the three categories listed above.

The consortia's plans for scoring

Here we are in somewhat unknown territory, and I find the prospects troubling. SBAC claims[11] that "The system—which includes both summative assessments for accountability purposes and optional interim assessments for instructional use—will use computer adaptive testing technologies to the greatest extent possible to provide meaningful feedback and actionable data that teachers and other educators can use to help students succeed." PARCC will "use technology throughout the design and implementation of the assessment system. The overall assessment system design will include a mix of constructed response items, performance-based tasks, and computer-enhanced, computer-scored items. The PARCC assessments will be administered via computer, and a combination of automated scoring and human scoring will be employed" (http://www.parcconline.org/parcc-assessment-design).

I have several concerns that have to do with computer-based "efficiency." The stated goal (in person, if not in print) of both consortia is to move toward the point where all assessments will be not only given on computers, but also completely computer-scored. Although advocates argue otherwise, I am far from convinced that the state of the art with regard to the automatic grading of "essay questions" in mathematics – especially those that employ diagrams and other mathematical representations – is anywhere near the point that student work on complex open response questions can be accurately assessed. (Indeed, computer entry can be a problem. Where one might draw a sketch and write some equations in a short amount of time, entering the same information into the computer could be a long, tedious, and distracting process.)

I have equally large reservations about the very notion of computer-adaptive scoring. Such scoring may not be too dangerous when the goal is to simply assign one score, and reporting on content and practices is not central. (Note that that should not be the case here.) But

[11] http://www.smarterbalanced.org/smarter-balanced-assessments/

worse, students who get off to a shaky start by giving the wrong answers to the first two problems on a test with computer-adaptive scoring may never have the opportunity to demonstrate what they know. The primary determinant of the “next” questions in computer-adaptive testing is item difficulty, the goal being to converge rapidly on a student score. This may be efficient, but it does not serve the needs of students or teachers in providing information about what students know and can do.

To put things bluntly, the primary concern of those who actually construct the exams must be the mathematics described in the CCSSM – not efficiency and not psychometric concerns such as reliability and validity. Those concerns are important, of course – but if they drive the construction of the assessments, they may well distort them.

Sample released Items.

I preface my comments by noting that it is early in the process, and a rather small number of sample items has been released by PARCC – so it is dangerous to extrapolate from what has been made available thus far. Readers should look at the collection of items – PARCC’s at < http://www.parcconline.org/samples/item-task-prototypes> and SBAC’s at <http://sampleitems.smarterbalanced.org/itempreview/sbac/index.htm> – and form their own opinions. My overall sense of the SBAC items is that, while the dynamic presentations in some of the items (e.g., item 43025) is superfluous, the content is reasonable and the use of the medium is on target and sometimes creative. Item 42960, for example (Figure 4), is straightforward and deals with relevant content. The computer-based format improves on the “matching” format used in many paper and pencil or computer tasks. The sample extended items call for student use of the practices, and the rubrics for scoring seem appropriately targeted. (One can always quibble about details, of course.)

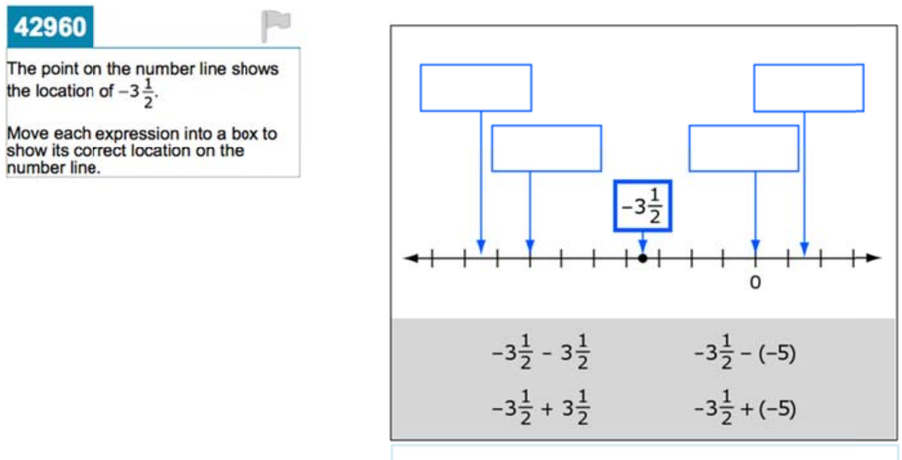


Figure 4. A computerized version of a "matching" problem.

PARCC givers readers less to examine, and some of what it offers is problematic. Consider, for example, the task in Figure 5.

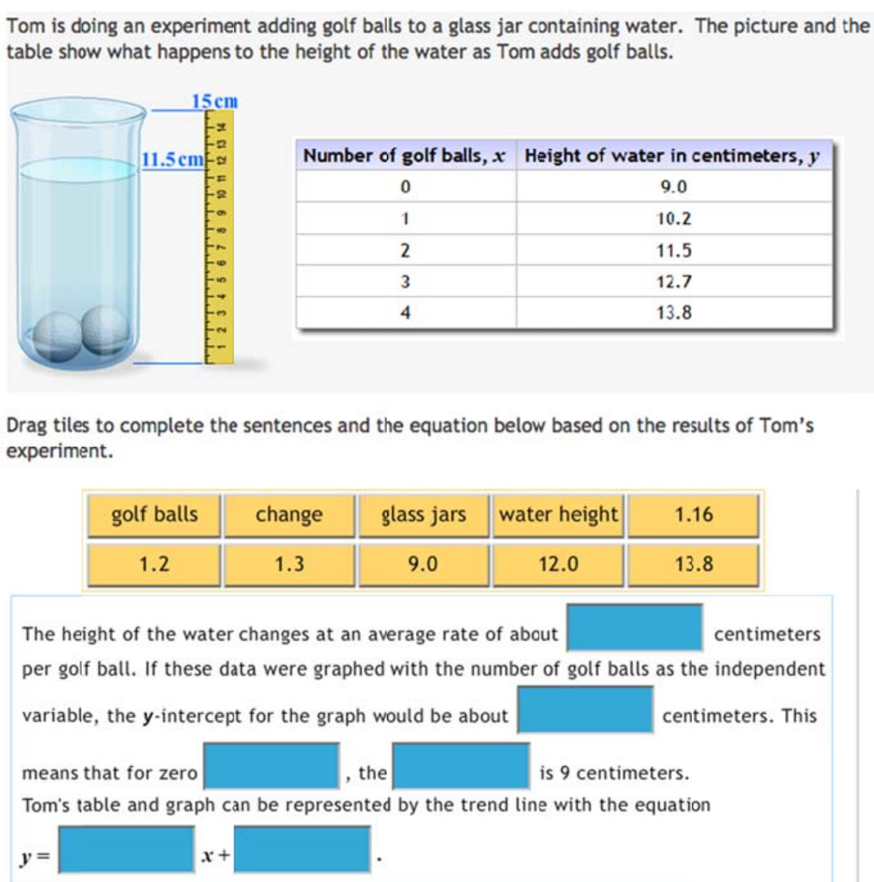


Figure 5. Part a of a PARCC modeling task.

Like the task in Figure 4, the task in Figure 5 asks students to drag tiles. Here, however, the use of computer technology in this task seems more of a distraction than a help.

In sum, there is cause for concern. Given that the tests will have such a powerful impact on mathematics instruction in the US, it is important to get things right.

Formative assessment

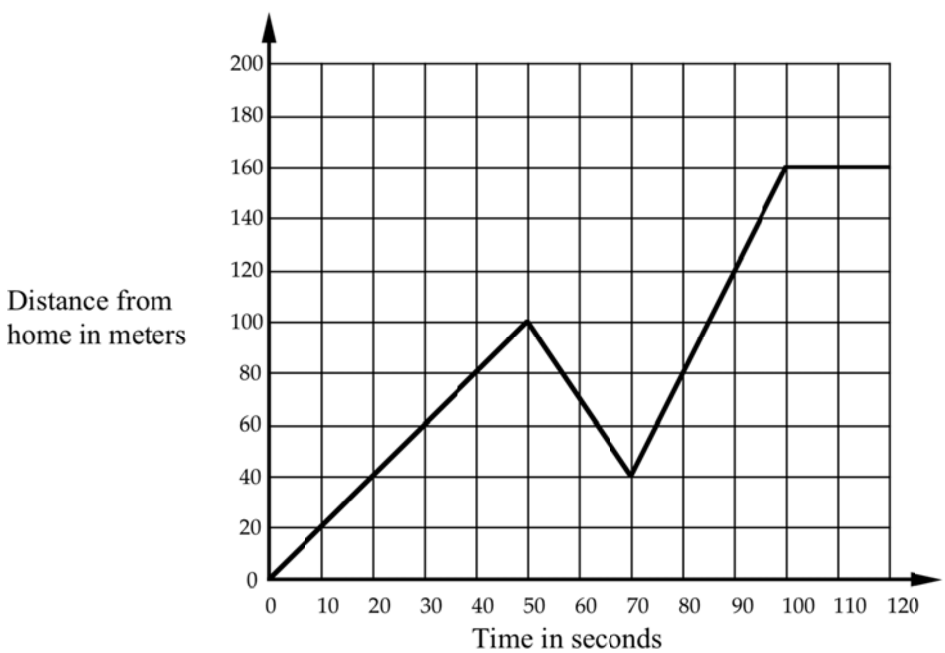
A major challenge facing teachers, especially those whose instructional focus has primarily been on procedural items such as the item exemplified in Figure 2, is to learn how to provide students with the skills and understandings required to address tasks like the one in Figure 3. Part of that challenge is learning to deal productively with student approaches – both correct and incorrect – as students grapple with complex tasks. A major approach to doing so is known as *formative assessment.* The purpose of formative assessments is not simply to show what students "know and can do" *after* instruction (that is the kind of summative assessment discussed above), but to reveal their current understandings so that the teacher can help the students improve. There is a large literature on formative assessment, but I will skim the surface of that literature by highlighting two main points. First, formative assessment is *not* summative assessment given frequently. As noted above, the purpose of formative assessment is to provide information about student understanding at a point when the teacher and students can act productively on that understanding. Second, the point of formative assessment is not to assign scores, but to highlight conceptual strengths and challenges faced by the students. In fact, scoring student work (even work that is heavily commented upon) increases the likelihood that the teacher's comments will go unread (Black & Wiliam, 1998).

The Mathematics Assessment Project (MAP), for which I am Principle Investigator, has been producing formative assessment lessons (FALs), whose purpose it is to support teachers in conducting formative assessments. As I write, 60 FALs (as well as a number of other assessment-related resources) are available for free on the MAP web site, <http://map.mathshell.org/materials/index.php>. To convey the flavor of the approach taken by the project, I briefly describe the FAL "Interpreting distance-time graphs," <http://map.mathshell.org/materials/lessons.php?taskid=208&subpage=concept>.

FALs begin with a diagnostic problem that the students work *before* the lesson, so that the teacher is provided information about the students' likely strengths and pitfalls. The diagnostic problem for "interpreting distance-time graphs" is given in Figure 6.

Journey to the Bus Stop

Every morning Tom walks along a straight road from his home to a bus stop, a distance of 160 meters. The graph shows his journey on one particular day.



1. Describe what may have happened.
   You should include details like how fast he walked.
2. Are all sections of the graph realistic? Fully explain your answer.

Figure 6. Diagnostic problem for "interpreting distance-time graphs."

The FAL lesson plan suggests that the teacher respond to the student work not by assigning scores, but instead by creating a set of questions that address the issues revealed by what the students have written. As support for the teacher, it identifies typical student misinterpretations and suggests questions that might push student thinking further. Common issues include (a) Student interprets the graph as a picture; (b) Student interprets graph as speed–time; (c) Student fails to mention distance or time; (d) Student fails to calculate and represent speed; (e) Student misinterprets the scale; and (f) Student adds little explanation as to why the graphs is realistic. A sample set of questions for issue (a) is given in Figure 7.

Common Issue	Suggested Questions and Prompts
Student interprets the graph as a picture For example: The student assumes that as the graph goes up and down, Tom's path is going up and down. Or: The student assumes that a straight line on a graph means that the motion is along a straight path. Or: The student thinks the negative slope means Tom has taken a detour.	• If a person walked in a circle around their home, what would the graph look like? • If a person walked at a steady speed up and down a hill, directly away from home, what would the graph look like? • In each section of his journey, is Tom's speed steady or is it changing? How do you know? • How can you figure out Tom's speed in each section of the journey?

Figure 7. A sample student issue and questions to explore it.

The goal is for the teacher to annotate the student work (individually if time permits, or by way of a list of "thought questions" for the class if not), so the students can engage more fully with the content. The full 90-minute lesson begins with a whole-class discussion of the problem in Figure 8. The students are asked to decide which of the stories A, B, and C corresponds to the distance-time graph that appears in the figure, and there is a whole-class discussion of the reasons students had for their choices. The result of this discussion is an annotated graph, which looks something like Figure 9.

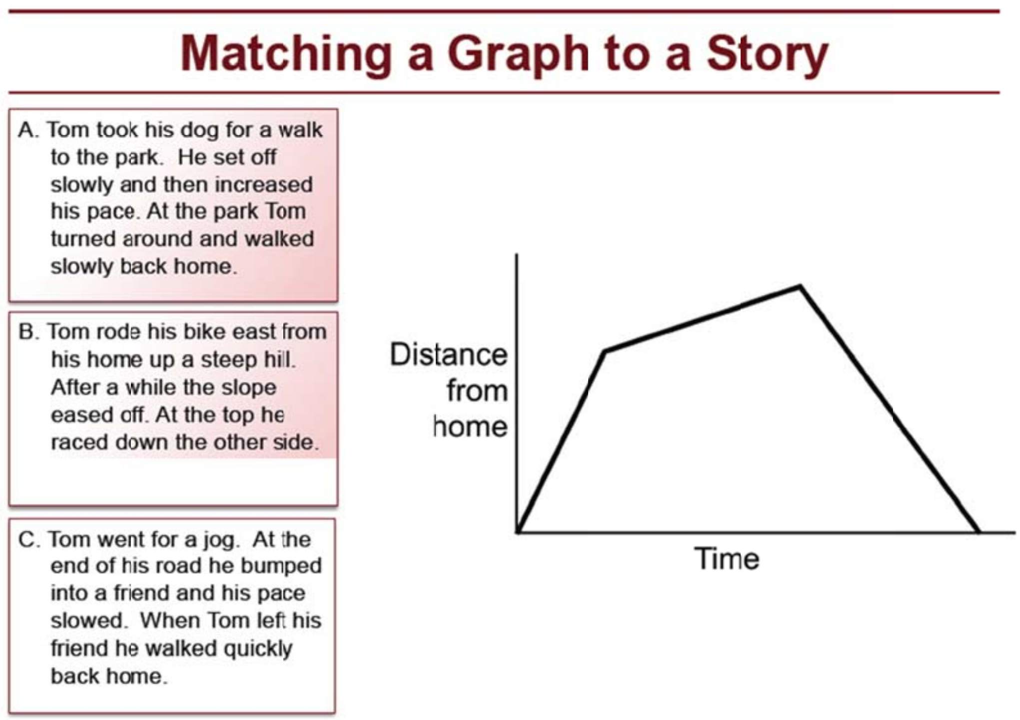


Figure 8. A distance-time graph question to start the lesson.

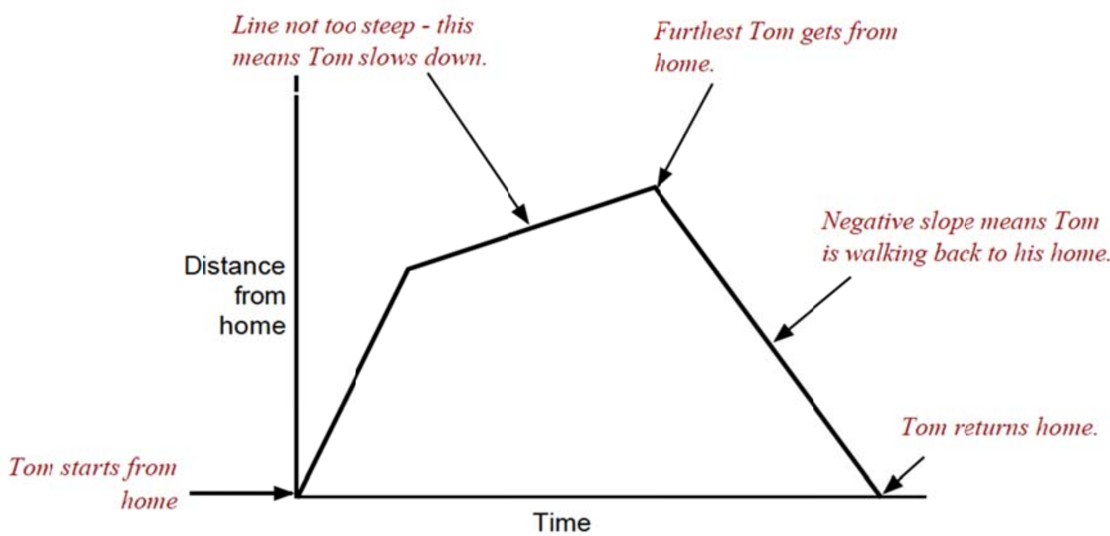


Figure 9. An annotated graph.

With this as backdrop, the main part of the lesson, a card matching exercise, begins. Students are given a set of ten distance-time graphs and ten stories. They are asked to work in small groups, matching the stories to the graphs. A sampling of the first four distance-time graphs is given in Figure 10.

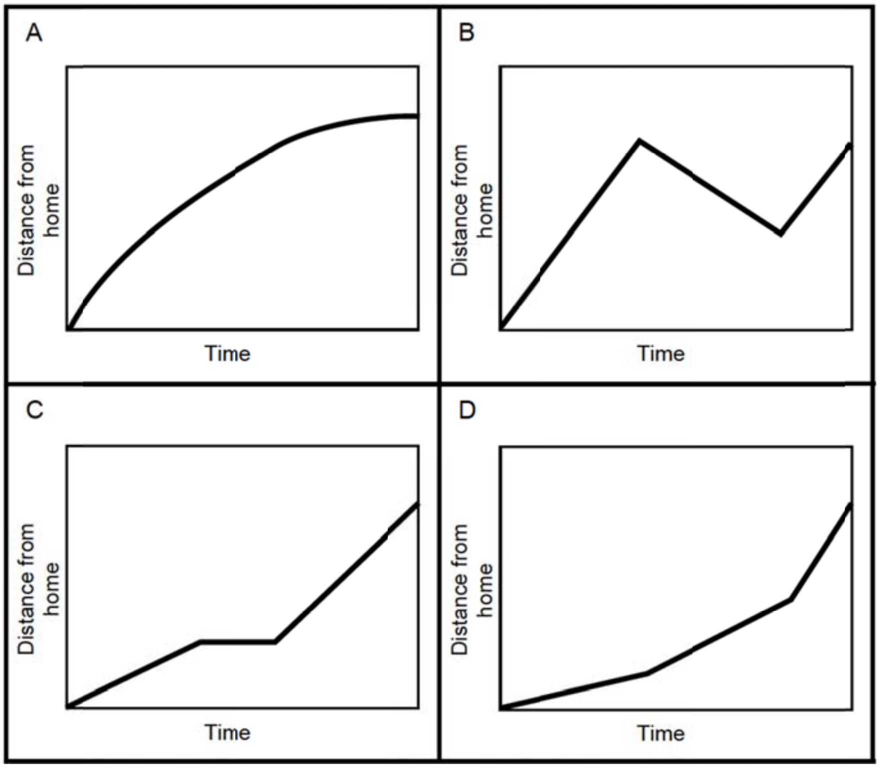


Figure 10. Sample distance-time graphs for the card sort.

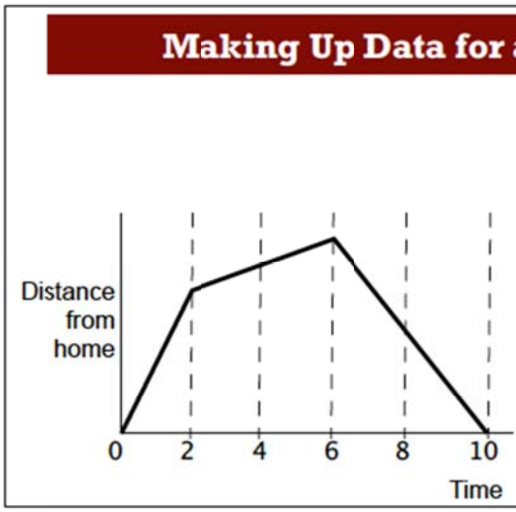
Four of the nine filled-out stories are shown in Figure 11. The tenth card says "make up your own story."

1 Tom ran from his home to the bus stop and waited. He realized that he had missed the bus so he walked home.	**2** Opposite Tom's home is a hill. Tom climbed slowly up the hill, walked across the top, and then ran quickly down the other side.
3 Tom skateboarded from his house, gradually building up speed. He slowed down to avoid some rough ground, but then speeded up again.	**4** Tom walked slowly along the road, stopped to look at his watch, realized he was late, and then started running.

Figure 11. Sample stories for the card sort.

As students work on the sorting task, they often encounter untenable situations – e.g., they have two incommensurate stories for the same graph, or two different graphs for the same story. This gives rise to heated conversations about *why* stories and graphs do or do not match.

At this point in the lesson, the teacher, who has been monitoring the discussions, starts a conversation about how to resolve the conflicts. He or she introduces the idea of building a table from a graph. Consider the graph in Figure 8, for example. One can assign a scale for time and distance to the graph, and then use selected points from the graph to generate a table. (See Figure 12, in which values for time has been assigned to the horizontal axis.) The completed table can then be uses to address questions such as, “Was Tom moving more rapidly in the first or second segment of the trip represented in the figure?”.



Time	Distance
0	
2	
4	
6	
8	
10	

Figure 12. Building a table from the graph.

Now that the students know how to construct tables from the graphs and to use them as mediating devices to check the stories they think are related to the graphs, they are given a third set of cards, which contains a collection of distance-time tables. Their task now is to use the tables to reconsider their graph-to-story pairings, and to put together a poster that features ten matching triples, each containing a story, graph, and table that are mutually consistent. The students share their posters, compare and contrast results as a group. The lesson ends with students being given time to revise their posters on the basis of what had been discussed during the whole class discussion.

As noted above, this lesson is one of 60 (out of a planned 100 such lessons) that appear on the MAP website, < http://map.mathshell.org/materials/index.php>. The hope is that such materials will be used by teachers when they reach the relevant units in their curricula.

Discussion

The United States is at a crossroads with regard to mathematics education, with assessment playing a major role as a potential lever for change. The Common Core State Standards in Mathematics (CCSSM) represents the natural evolution of mathematical standards, dating back to the 1989 NCTM *Standards.* The potential for significant change comes with (a) the adoption of the CCSSM by 45 of the 50 states and three territories; (b) the fact that states that have aligned with the CCSSM will be using one of only two assessments (one produced by the PARCC consortium, one by SBAC) to assess student proficiency in mathematics. Condition (a) suggests that we will have, for the first time in the US, a de facto national curriculum. Condition (b) suggests that the two current assessments, because of the high stakes involved, will play a fundamental role in shaping how that curriculum comes to life in American classrooms. If the assessments focus on the mathematical values intended in the CCSSM, there is great potential for assessment-driven progress; but if the assessments pervert the mathematical intentions of the CCSSM writers for reasons of cost, ease in scoring, or psychometric considerations[12], the results can be disastrous. The stakes are indeed high, for (at minimum) the next decade of American mathematics instruction.

[12] It is absolutely essential for the mathematical integrity of the standards to drive the test construction process, with psychometric considerations then taken into account, rather than – as is typical in test construction – the other way around.

The right assessments can orient the system in the right directions, but even so, there are issues of system capacity. Teaching for the kinds of content understandings and mathematical practices described in the CCSSM is hard. Generally speaking, teacher preparation programs have not had the time or resources to help teachers become proficient at formative assessment; nor does the current generation of texts provide teachers with adequate support. Formative assessment, well done, can support teachers in building rich mathematical classroom environments. It is our hope that the kinds of formative assessment lessons (FALs) described in this paper will help to provide such support.

References

Barnes, M., Clarke, D., and Stephens, M. (2000). Assessment: the engine of systemic curricular reform? *Journal of Curriculum Studies*, 32 (5), 623-650.

Bell, A., & Burkhardt, H. (2001). Evidence on the influence of assessment on teachers' classroom practices. Downloaded October 31, 2001, from <http://www.toolkitforchange.org/toolkit/view.php?obj=1034&menu=i>.

Black, P., & Wiliam, D. (1998). *Inside the black box: raising standards through classroom assessment.* London: King's College London School of Education.

California State Department of Education. Standardized Testing and Reporting - STAR Sample Test Questions. Downloadable at < http://starsamplequestions.org/starRTQ/search.jsp>.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics.* Downloaded July 9, 2010, from http://www.corestandards.org/the-standards

National Commission on Excellence in Education. (1983). *A nation at risk.* Washington, DC: U. S. Government Printing Office.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington DC: National Academy Press.

Partnership for Assessment of Readiness of College and Careers. (2012a). PARCC model content frameworks: Mathematics grades 3–11, version 2.0, august 31, 2012 *(revised).* Downloaded September 23, 2012 from <http://www.parcconline.org/parcc-model-content-

frameworks>.

Partnership for Assessment of Readiness of College and Careers. (2012b). PPT_Messaging for PARCC Math Sample Items_FINAL_revised_081212_1. Downloaded from <http://www.parcconline.org/samples/item-task-prototypes> on September 23, 2012.

Smarter Balanced Assessment Consortium. (2012). Content Specifications for the Summative assessment of the *Common Core State Standards for Mathematics.* Downloaded March 20, 2012, from <http://www.smarterbalanced.org/smarter-balanced-assessments/>.

Swan, M., and the Shell Centre Team (1985). *The language of functions and graphs.* Nottingham, England: The Shell Centre of Mathematics Education.