SERP Strategic Education Research Partnership

Aligning Science Standards with Classroom Practice

M. Suzanne Donovan, SERP Institute

May 2009

SERP Publications, White Paper | www.serpinstitute.org

Aligning Science Standards with Classroom Practice
White Paper: May, 2009
Suzanne Donovan, SERP Institute

K-12 science education in the United States varies enormously not only from state to state and school district to school district, but also from school to school and from classroom to classroom within a school. State Standards and accountability assessments have become the primary mechanism for ensuring that, amidst the wide variation in practice, core content is being taught. The purpose of standards in a highly decentralized system is to create a baseline of shared expectations that can shape textbook content, focus teachers' instructional plans on important ideas, and galvanize those who provide professional development and teaching tools to focus their efforts on critical content. The assessments aligned to the standards provide a common metric that is critical to introducing accountability into a decentralized system.

Current standards, however, miss the mark. The standard setting committees convened to produce them by design included members who would collectively bring a broad range of expertise in all aspects of science. But when every committee member advocates for the arguably important content in his or her domain, the emerging set of standards is far too broad. The "mile wide-inch deep" problem in science education in the United States was reinforced and codified.

When standards are defined too broadly they wreak havoc: Textbooks that attempt to cover all standards include far more material than can possibly be taught in a single academic year, leaving teachers –many of whom have little preparation in science--with little or no guidance regarding the material that is critical and should be treated more fully, and material that can be set aside with less consequence. Inquiry and other forms of in-depth instruction gets crowded out by test preparation. Compelling visions of practice are moved off the table with a widely acknowledged excuse: *the standards require that we cover too much material to spend time going in depth on any one thing*.

The problem with existing science standards is broadly understood, and there are efforts under way to address the shortcomings. These include a commitment to identify "anchor standards,"[1] or standards that are "fewer, clearer, and higher."[2] These standards would privilege the ideas with the greatest scientific importance, giving permission to spend more instructional time on these and less on others. The hope is that with better standards we will see better science instruction.

Because breadth over depth is so pervasive a problem, we should expect this second effort at standard setting to represent an important improvement. And if the only problem were that the standards were too broad, replicating the previous standards setting process with a deliberate goal of producing narrower standards would represent an adequate response. But there is a second problem: the standards setting process in uninformed by the realities of the classroom. How much time is devoted to science

---

[1] National Science Teacher's Association, Science Anchors, http://scienceanchors.nsta.org/Default.aspx

[2] THE OPPORTUNITY EQUATION: Transforming Mathematics and Science Education for Citizenship and the Global Economy Report of the Carnegie Corporation of New York-Institute for Advanced Study Commission on Mathematics and Science Education. June, 2009.

instruction in a school year? How much time needs to be spent on the most important science content to ensure mastery? What related concepts need to be included in the standards if the most critical ideas are to be grasped by students? How can we assess all of these important standards in ways that are informative to both policy-makers and to teachers? These are empirical questions, but the standards process is not tied to an evidence-based process that shapes the outcome. Given that the very purpose of setting standards is to drive good instruction, this is a critical flaw.

The current effort to revise the science standards would be significantly strengthened if the plan of action were to be made more ambitious in two integrally linked ways: 1) by disciplining the standards-setting process to reflect the realities of the classroom; and 2) by putting in place a parallel program of research and development that generates the instructional tools (including formative assessments and instructional activities) that can support mastery of the material covered in the standards. The case for each is made below.

<u>Linking standards to classroom instruction</u>

The number and breadth of core standards should not be determined solely by a group of disciplinary experts who spend their professional lives in universities. Nor is adding a few practitioners –who are often silenced by their unfamiliarity or discomfort with the culture of discourse among academics—an adequate solution. While disciplinary experts have much to offer, their ideas should interact with evidence about what can be effectively taught in science classrooms in typical (or what we can realistically hope to achieve as typical) schools. Understanding the content for oneself is weakly linked to understanding what it takes to teach that content to elementary and secondary students. And it is unrelated entirely to an understanding of how many hours are available for science instruction once holidays, field trips, and assemblies have been taken into account.

Moreover, which specific ideas are included in the standards should be shaped by empirical evidence of the related concepts that students must master if they are to understand the core ideas. Experts often have "blind spots" because their thorough knowledge of the field can lead them to overlook foundational ideas that they take as given, but which many students do not know or fully understand, and which therefore will claim classroom hours to teach. Linking the standards setting process to an empirical exploration of student learning is likely to produce standards that support classroom instruction much more effectively. This, after all, is the holy grail of the standard setting movement. And creating assessments tied to these standards is key to making the advances we envision: to communicate what the standards mean for teachers, to allow policy-makers to have confidence in the results, and to "make visible" the learning for all involved.

<u>Instructional tools to support standards-based instruction.</u>

If we knew how much content teachers could realistically cover in a school year, revising the standards setting process to include that information would be straightforward. But we do not know. The results of the last round of standard setting produced a partial answer: far less than the standards committees assumed. The U.S. has not yet invested in the classroom based laboratories or field sites where researchers work alongside practitioners to build the knowledge base for professional practice. This helps explain why the first round of standards creation was so ineffective; the committees did not choose to ignore what is known about the realities of classroom practice, there simply was not a strong knowledge base on which to draw. To do better next time, we need to build that knowledge base.

But there is one more layer of complexity: even if there is broad agreement that the standards will target average students (neither remedial nor honors), the answer to how much can be covered in a year is almost certainly this: it depends on what we mean by "covered," and on the effectiveness of the teacher. For a more specific answer, we will need to create and test compelling, concrete examples of what it would mean to do science instruction well, to adapt these examples for use in assessments, and then build the tools that will allow a broad range of teachers to engage in effective instruction. To achieve that end we will need to bring together interdisciplinary teams of researchers not just to deliver wisdom, but to work with practitioners to build these supports for instructional practice. A world in which standard setting both informs, and is informed by, this type of R&D holds promise for real progress in science teaching and learning.

The problem, of course, is that an ambitious program of research and development will take time, and there is little sense in holding improvements that are within reach (narrowing the standards to a smaller, more important set) hostage to building a program of research and development in classroom settings. Why not move forward with the standard setting process, and do the rest later?

The political will to set any change in motion is mustered only with a great deal of effort. Political "windows of opportunity" that open rarely remain open for long. Once this second phase of standard setting is complete, the window is likely to close again; we will be done with standards improvement until evidence of unintended consequences and unanticipated failures mount over years, eventually forcing (with intense effort) the window open once again.

Furthermore, once standards are set, the enthusiasm of school district professionals to engage in the work required to feed a better outcome next time is likely to be undermined by a sense of fatalism. When the standards have been set through a process that is entirely top-down, teachers and administrators feel disempowered. We are far more likely to be successful in supporting real change if we seize the opportunity now to redefine the problem we are trying to solve. Rather than seeing the end product as the new set of standards, we could set out to put in place a system for ongoing improvement, in which standards, assessments, and instruction inform each other regularly and continuously. This would mean defining standards setting as a periodic process in which an ongoing program of research and development will feed planned, empirically-based improvements.

Managing Complexity

There are two mistakes to be avoided, each of which must be taken seriously. The first would be to err, as argued above, by once again designing a process that is blind to the realities of the classrooms it is attempting to influence. The second would be to err in the opposite direction: by weighing down the standards setting process excessively, making it too complex and therefore politically unappealing. Both errors could be avoided if the standards setting process and the R&D process were on parallel tracks, heading toward the same destination but moving forward at different speeds. The key would be to have them stop at the same stations along the way.

In the "working period" during which standards are being negotiated, an initial investment could be made in the program of R&D that will inform the standards setting process over time. Critical members of the R&D team could be at the standards-setting table, bringing to bear evidence that has already been gathered from school-based work, and that is being generated during the working period. The standards train would move far faster along its track, but when the round of standards setting is complete, it would rest for some time (possibly 5 years) at the next station. The school district/research partnerships involved in the R&D effort would continue to move more slowly, gathering information about how to support effective instruction of the standards content. When it reaches the first station, the R&D train will be loaded with evidence. Are there too many standards? Or are some critical standards missing? Do the standards miss critical links across years? Do some things need to be put into the 7th grade standards in order to teach effectively to the 9th grade standards?

If we acknowledge the complexity of effective teaching, the idea that standards-setting should be an ongoing process that is fed by evidence makes far more sense than envisioning a point in time activity that can be called complete when a deadline for a working set of standards is met. While most of the players in education policy and practice know this to be the case, the vastly different timelines for policy (standards setting) and research and development (how does/can it play out in schools) has driven those with primary commitments in one or the other of these arenas in opposite directions. Establishing parallel tracks for these efforts designed to stop at common stations will allow both groups to begin moving in the same direction.

A Place to Start

Constructing the R&D program described presents a major design challenge. How can we envision an R&D effort that would produce the desired results? In its San Francisco field site, a collaboration formed by the Strategic Education Research Partnership (SERP) provides a glimpse of how such a program might look. The problem posed by the San Francisco Unified School District leaders to the SERP team was focused on middle school students, and was roughly as follows:

San Francisco had for many years worked with partner organizations on strengthening science instruction. But the standards and accountability movement intervened in ways that set the science instruction clock back. Accountability requirements meant that less time was available for science education in many schools (because more was required for reading and math), and that coverage of more science content was expected. Adoption committees for middle and high school science chose traditional texts because they covered all the standards and provided a vehicle for moving through a great deal of material quickly. Inquiry instruction, or any other form of in-depth treatment of content--became less common because "coverage" became the primary goal. However the textbooks were so broad that a good deal of the content could not possibly be covered, even if all in-depth instruction was abandoned. What actually got taught differed greatly from one classroom to the next.

District leaders were in search of a realistic solution. How can students be given the opportunity to experience the process of doing science by engaging in experimentation without putting teachers in the untenable position of sending their students off to high stakes tests never having seen much of the content? District leaders have a good sense of what quality science instruction looks like. They consider it their jobs to support good science instruction in the district's classrooms. But they, too, feel at the mercy of the standards.

In the context of the SERP collaboration (which includes district teachers and administrators, and researchers from the University of California, Berkeley, Stanford University, University of California, San Francisco, the Lawrence Hall of Science, San Francisco State University, and the Exploratorium) the problem was shaped as follows:

Teaching science content in depth takes more time and effort. The breadth required by standards drove teachers away from teaching content in depth; but the shift happened only too easily for most teachers because teaching in depth is so effortful, and because it requires more knowledge and teaching expertise. Asking students to read and memorize what is in a textbook is straightforward. Setting up experiments and supporting students to learn from those experiments is in contrast much more challenging. To make the task of teaching some content in depth feasible, the targeted concepts must be relatively few, and the supports for effective teaching relatively extensive.

As a place to start, the science collaboration decided to break the school year into "marking periods." The working assumption of the group was that teachers could realistically go in depth on one major idea per marking period. A SERP "working group" of experienced science teachers and researchers with expertise in science, science education, and assessment are engaged in a process (led by the U.C. Berkeley BEAR Center) that includes the following:

a) identifying the critical content (4 or 5 major concepts per year) for in-depth instruction
b) developing progress maps for each of those core ideas that capture the thinking students are likely to display as they move from an everyday, naive understanding of the concept toward an accurate scientific understanding. The maps are drafted using the research literature and the knowledge of expert teachers, then put to the reality test in the teachers' classrooms.
c) Developing formative assessment tasks that allow teachers to see where student understanding falls on that progress map
d) Developing (or adopting existing) well designed, fully supported instructional activities (including labs or inquiry activities) that can move students effectively from a lower to a higher level of understanding

While the description is linear, the work is not. The process of developing and experimenting with a progress map in classroom settings identifies weaknesses in students' understanding of related concepts. Misconceptions are uncovered that require the development of another progress map inside the original progress map. Formative assessments must be developed that both capture the science content well, and have desirable psychometric properties. The process reveals the underlying complexity of the challenge. It is time consuming, and the roadmap is being created as the work progresses. But working through that complexity promises to provide the foundational knowledge for describing the core ideas that belong in the standards, *and* a first pass at the critical knowledge for effective instruction in that content.

<u>What would it take</u>?

The process just described, with its focus on student learning and classroom instruction, stands in stark contrast to the typical standards creation process that is rooted entirely in scientists understanding of their field. It has at its core a set of assumptions regarding the ideas that are powerful enough to serve as anchors for student understanding of an area of science. This R&D process could feed the national effort if the starting point for the development of tools were the ideas privileged by the standards. Linking the two efforts would strengthen both.

To support a convergence of the two approaches to defining core content, three things would need to happen:

a) The states that are now setting out to develop common standards would need to redefine the effort as an ongoing one, and establish the expectation that the revision process will occur at some expected time—such as at 5 year intervals.

b) A philanthropic organization or a government agency (or a combination of the two) would need to invest in an expanded and well-coordinated R&D effort. While the work under way in San Francisco can serve as a model, the scope, scale, and pace of the effort tied to the standards would need to be considerably greater.

c) Key participants in the R&D would need to be at the standards setting table in order to serve two roles: 1) contributing evidence from the program of work that can shape standards settings, and 2) feeding the proposed standards back into the R&D work in order to generate empirical evidence regarding their use in practice.

There is a good deal of relevant work that has already been done, or that is under way (much of it supported by the National Science Foundation). But that work needs to be organized for the purpose at hand, and systematically built upon.

<u>A Role for SERP?</u>

The SERP Institute would be pleased to support an effort to link standards setting to classroom practice in any way, large or small, that would be productive. SERP has for almost five years been successfully bringing distinguished scholars from some of the best universities in the nation into school settings to work on problems of practice. The SERP process under way in San Francisco could serve as a starting point for contributing to more effective standards settings processes. But a far broader community of researchers would need to be recruited to the effort, as would a set of school districts in a variety of states and district settings.

Should a major research and development effort get traction, SERP would be pleased to coordinate and/or support the expanded effort described above. Because the SERP organization has no allegiance to other institutions or political entities, it is well situated to broker the partnerships across organizations, drawing together people in disparate universities and research institutions. Its track record in recruiting distinguished, interdisciplinary teams, and in working effectively in school district contexts, would be an asset.