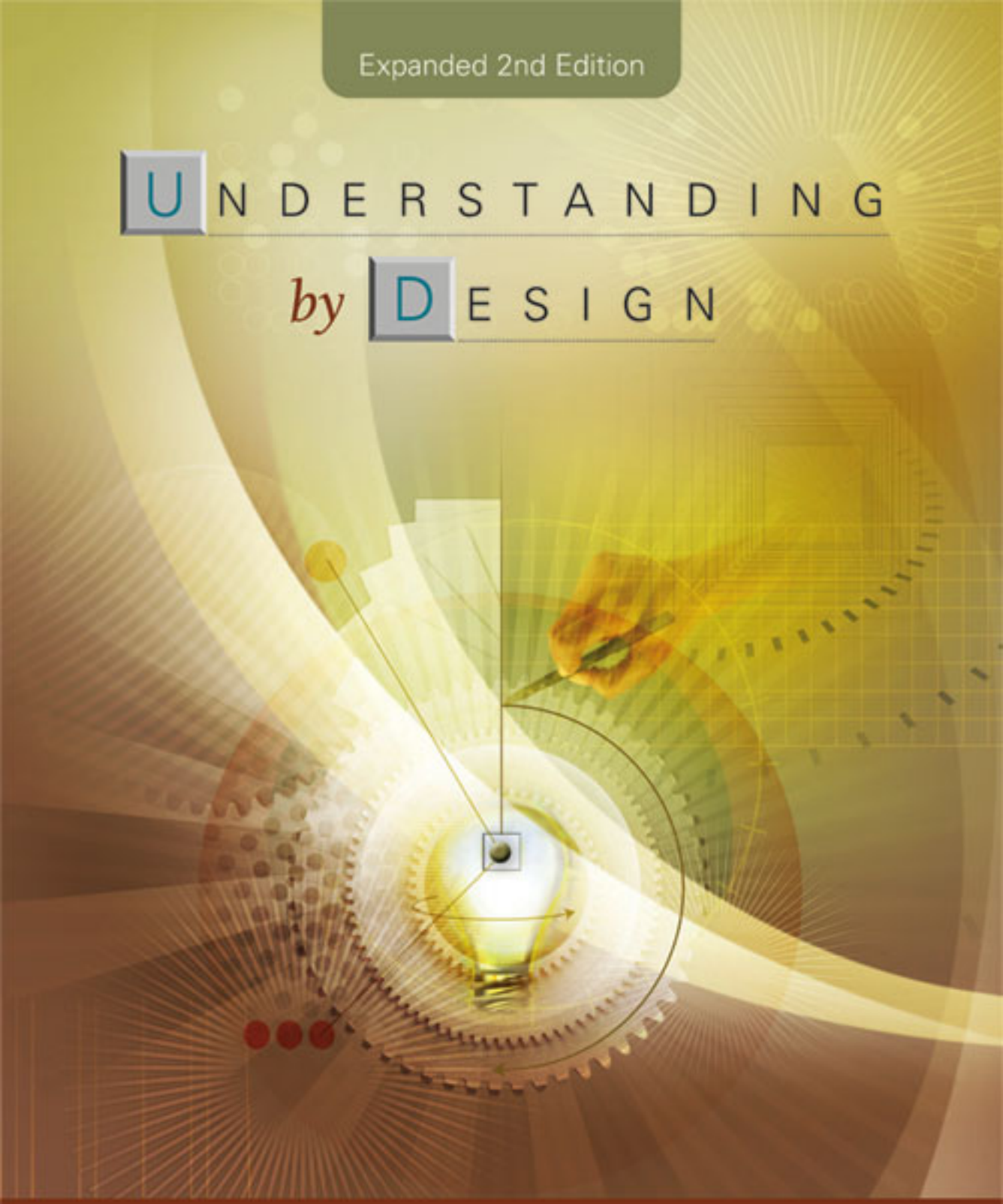
Expanded 2nd Edition

UNDERSTANDING *by* DESIGN



GRANT WIGGINS AND JAY MCTIGHE

Chapter 13

"Yes, but . . ."

> The work is difficult and requires constant revision. It is particularly difficult for teachers who have to "unlearn" their prior practice.
> —Mark Wise, Social Studies Supervisor, West Windsor-Plainsboro, NJ

> Now consider what their release and healing from bonds and folly might be like. . . . Take a man who is released and suddenly compelled to stand up, to turn his neck round and look toward the light . . . who, moreover, in doing all this is in pain, and because he is dazzled, is unable to make out the shadows he knew before. . . . And if he were compelled to look at the light itself, would his eyes hurt and he would flee . . . and if someone dragged him by force along the rough ascent, wouldn't he be distressed and annoyed? He would.
> —Plato, *The Republic,* c. 360 B.C.E.

In this book, we have set forth a vision of and pathway toward meaningful curricular, assessment, and instructional reform, all carefully designed around planning for understanding. We understand that our reform vision is neither highly original nor very radical. It parallels the vision of many educators, researchers, and reformers in the past decades.

Nonetheless, whenever reform ideas are proffered, it is common to hear a chorus of "Yes, but . . ." from well-intentioned teachers and administrators. The proposed reforms are damned with faint praise and undercut by the rejoinder that these fine ideas cannot work in today's world of state standards and high-stakes testing. Some reformers remain adamant that good pedagogy and state standards and testing are inherently incompatible; many educators worry that there might not be a research base to support our arguments, no matter how commonsensical.[1]

We empathize with these laments and the concerns they are based on, given the accountability pressures facing educators. Yet many of the recurring arguments are based on misunderstandings about learning, assessment, standardized testing, teaching for understanding of big ideas, and the relationship between local pedagogy and state standards. In this chapter, we provide the

arguments and research in support of our views while examining three key misconceptions that often hold back or interfere with comprehensive reform. We explain why each *is* a misconception by "unpacking" the implicit and questionable assumptions underlying the "yes, but . . ." concerns, and offer a friendly but firm rebuttal.

The misconceptions we address:

"Yes, but . . . we have to teach to the state and national tests."

"Yes, but . . . we have too much content to cover."

"Yes, but . . . the needed curriculum and assessment work is hard and I simply don't have the time to do it well."

Misconception 1: "Yes, but . . . we have to teach to the test."

State, provincial, or national content standards and concomitant testing programs have emerged worldwide with the intention of focusing local curriculum and instruction on boosting student achievement by holding schools accountable for results. Ironically, the key lever in this standards-based reform strategy—the use of high-stakes external tests—has unwittingly provided teachers with a rationalization for avoiding or minimizing the need to *teach well*, that is, to teach for in-depth understanding.

For many educators, instruction and assessing for understanding are viewed as incompatible with state mandates and standardized tests. Though research is rarely offered to support this oft-heard claim, the speaker clearly implies that school faculties are stuck teaching to the test—against their will. They *would* teach for understanding, if they could. The implicit assumption is key: The only way to safeguard or raise test scores is to "cover" those things that are tested and "practice" the test format (typically selected-response or brief constructed-response) by having local assessment mimic state assessment. By implication, there is no time for in-depth and engaging instruction that focuses on developing and deepening students' understanding of big ideas; nor is there time for performance assessment.

This opinion is so widely held that many readers may be thinking that *we* are the ones harboring the misunderstanding (or myopia or naïveté) about the real world of education. Isn't it a fact that we have to teach to the test and leave aside more higher-order, big-idea-focused, and performance-based approaches? Many certainly think, say, and act accordingly. Although we are obligated to teach to content standards, it does not follow that the best way to meet those standards is to mimic the format of the state test in all local testing and haphazardly cover all prescribed content through superficial and scattered teaching.

To more clearly show why the common complaint and reluctant solution are based on a misunderstanding, consider a rephrasing of the reason given for focusing on test items at the expense of depth. The speaker asks us to

believe that the only way to *raise* test scores is to teach *worse*. That is not how the speaker usually puts it, of course, but that is what the argument amounts to. "I would love to teach for understanding, but I just can't; it won't pay. I'm better off teaching discrete facts and skills, just the way they are tested" is what the first "yes, but . . ." response really means.

Just putting it this way should cause a raised eyebrow or two. Is it *really* either/or? Must we avoid effective and engaging forms of instruction to *raise* test scores? Is more passive, fragmented, and superficial teaching *more* likely to maximize student interest and performance? We think this theory is incorrect, based on a misunderstanding about how testing works.

The parallel with the doctor's physical

To begin to uncover the flaw in this reasoning, consider an analogy. Once a year, we go to the doctor for a physical exam. No one particularly relishes the thought of such an exam, but we go with the understanding that it is in our long-term interest to get an objective (yet superficial) measure of our health. In fact, it is more like an audit because the nurse and lab technicians perform a few tests in a short span of time (such as blood pressure, pulse, temperature, blood work for cholesterol). The physical is a small sample of tests, yielding a few useful health status indicators. Its validity and value stem from the fact that the results *suggest* our state of health, not because the physical *defines* healthfulness. We experience a relatively quick and unintrusive physical exam so that various indicators can be examined for signs of trouble demanding further scrutiny.

Now suppose we are terribly concerned about the final numbers (e.g., weight or blood pressure) and that the numbers ultimately link to our personal health insurance costs. What we might do, in our panicky state prior to each annual physical, would be to "practice" for the test—focus all our energy on the physical exam (as opposed to what its indicators suggest). If our doctors knew of our actions, their response would surely be something like this "Whoa! You have it backward. The best way to 'pass' your physical is to live a healthful life on a regular basis—exercising, watching weight, lowering intake of fats, eating more fiber, getting sufficient sleep, and avoiding tobacco. You're fixating on the indicator instead of on the causes of good results."

Why? None of the elements of true healthfulness—your diet, your fitness regimen—are tested directly in the physical; doctors audit your health *indirectly* through factors including blood pressure, weight, skin tone, and color. Thus, "normal blood pressure" and "normal weight" are only indicators of overall wellness and fitness, not to be confused with overall health. The physical exam involves assessing a few quick, usually accurate indicators. So to confuse the indicator with the thing itself is poor policy. The more that you concentrate only on your weight, for example, to the exclusion of everything else in your daily regimen, the less likely it is that you will be healthy in the long run.

Like the doctor, state education agencies give schools a "checkup" once a year by viewing indirect evidence—state tests—of student intellectual health. A test, like the physical exam, is an audit related to the state standards. Like the physical, the state test provides indirect indicators about our health. Test items indirectly assess the quality of our "daily regimen" in the same way that a look at blood pressure and weight are proxies for the daily "tests" of real fitness and wellness.

We *can* get some good information about the rigor of our regimen from quick-and-dirty indicators. Any good test—whether in the school or the exam room—need not involve the core performance we should be engaged in daily. For schools, it only matters that the indicators yield valid inferences to the standards. That is the nature of test validity, as we saw in previous chapters—establishing a link between one set of easy-to-obtain indicators with a related set of complex and desired results.[2]

It would be thought silly to practice for the physical exam as a way to be healthy. But this error is precisely what we see in too many schools all over North America. Local educators, fearful of results, are focusing on the indicators and not the causes of happy results.

Please understand that this explanation does not constitute an endorsement of any specific test question or current state practice in which we rely heavily on one-shot external testing, often done in secret with tests improperly vetted. In fact, we feel strongly that state agencies and policy makers bear a responsibility for allowing the confusion about the relationship between local practice and state tests to persist by not making local assessments part of a comprehensive state accountability system, and by not making more of an effort to design more transparent accountability (such as through release of all tests and results once the test has been administered, for the sake of both feedback and fairness).

What matters for local reform is that we take to heart the point of the analogy: *We* are responsible for wellness, not the state. The state's job is to audit—just as the physical exam does—not to provide the daily regimen we should engage in at home. Indeed, the state could not possibly assess everything of value in an authentic way, even if we all wanted it, because of excessive costs and the desire to limit the intrusions of external testing. Doctors have a similar problem—requiring every patient to come in for a multiday comprehensive fitness program and workup at a medical lab would be excessively time-consuming and costly (never mind the unlikelihood of getting our insurers to foot the bill). So, in the absence of data to show that the indicators yield invalid inferences, the task is to focus on local rigor, not test prep.

The misunderstanding about what is cause and what is effect in performance gains may well be related to misunderstandings about the "face validity" of tests, as test-makers term it. Educators might look at both test format and content and conclude that the test neither rewards teaching for understanding nor performance-based local assessment. That view, while understandable, is mistaken. Validity is about the empirical link between test results, the

objectives tested, and local practice. That is why tests that appear inauthentic can yield valid inferences (e.g., vocabulary tests are often good predictors of academic success) if designed properly, and why some performance-based projects yield poor results (since the projects often end up unrelated to state standards, as in the diorama example discussed in Chapter 9). Making matters worse, many teachers then erroneously infer that instructional practice is somehow dictated by test format, so they teach a *random and superficial survey of content*—making it far less likely that student learning will be engaging and effective.

To invoke a different analogy to explain the error in logic, state standards are like building codes; local instructional design is our architecture. The goal of architectural design is not to meet building and zoning codes in a slavish fashion. The goal is to design something that is practical, pleasing, and stylish—while meeting building and zoning codes.

In fact, the situation regarding education is far better than many assume. Most state standards stress the importance of in-depth understanding and mastery of key complex performances and genres in which knowledge, skill, and understanding are revealed. Understanding by Design (and other programs and reform approaches) provides a way in which a focus on big ideas, robust assessment, and a focused and coherent learning plan makes it likely that state standards are addressed and met.

Research base

The best news is there is an empirical basis to this logical argument. In the mid-1990s, Newmann (1996) and others conducted a study of restructured schools at the elementary, middle, and high school levels. This ambitious study measured how well 24 restructured schools implemented authentic pedagogy and authentic academic performance approaches in mathematics and social studies, and whether schools with high levels of authentic pedagogy and academic performance significantly increased achievement over those that measured at low levels. Authentic pedagogy and performance were measured by a set of standards that included higher-order thinking, deep-knowledge approaches, and connections to the world beyond the classroom. Selected classes were observed four times during the school year in each school. The researchers observed 504 lessons and analyzed 234 assessment tasks. They also analyzed student work.

Students in classrooms with high and low levels of authentic pedagogy and performance were compared, and the results were striking. Students in classrooms with high levels of authentic pedagogy and performance were helped substantially whether they were high- or low-achieving students. Another significant finding was that the inequalities between high- and low-performing students were greatly decreased when normally low-performing students used authentic pedagogy and performance strategies and assessments.

The study provides strong evidence that authentic pedagogy and assessments pay off in improved academic achievement for all students, but especially for low-performing students. This research supports the Understanding by Design approach, which emphasizes the use of authentic performance assessments and pedagogy that promotes a focus on deep knowledge and understanding, and active and reflective teaching and learning.

Two recent studies of factors influencing student achievement were conducted in Chicago public schools through the Consortium on Chicago School Research. In the first study, Smith, Lee, and Newmann (2001) focused on the link between different forms of instruction and learning in elementary schools. Test scores from more than 100,000 students in grades 2–8 and surveys from more than 5,000 teachers in 384 Chicago elementary schools were examined. The results provide strong empirical support that the nature of the instructional approach teachers use influences how much students learn in reading and mathematics. More specifically, the study found clear and consistent evidence that interactive teaching methods were associated with more learning in both subjects.

For the purposes of the study, Smith, Lee, and Newmann characterized interactive instruction as follows:

> *The teacher's role is primarily one of guide or coach. Teachers using this form of instruction create situations in which students . . . ask questions, develop strategies for solving problems, and communicate with one another. . . . Students are often expected to explain their answers and discuss how they arrived at their conclusions. These teachers usually assess students' mastery of knowledge through discussions, projects, or tests that demand explanation and extended writing. Besides content mastery, the process of developing the answer is also viewed as important in assessing the quality of the students' work.*
>
> *In classrooms that emphasize interactive instruction, students discuss ideas and answers by talking, and sometimes arguing, with each other and with the teacher. Students work on applications or interpretations of the material to develop new or deeper understandings of a given topic. Such assignments may take several days to complete. Students in interactive classrooms are often encouraged to choose the questions or topics they wish to study within an instructional unit designed by the teacher. Different students may be working on different tasks during the same class period. (p. 12)*

The type of instruction found to enhance student achievement parallels methods advocated by Understanding by Design for developing and assessing student understanding. Smith, Lee, and Newmann (2001) summarize their results as follows:

> *The positive effects of interactive teaching should allay fears that it is detrimental to student achievement of basic skills in reading and mathematics. Conversely, the findings call into serious question the assumption that low-achieving, economically disadvantaged students are best served by*

emphasizing didactic methods and review. Our results suggest precisely the opposite: to elevate mastery of basic skills, interactive instruction should be increased and the use of didactic instruction and review moderated. (p. 33)

A related study (Newmann, Bryk, & Nagaoka, 2001) examined the relationship of the nature of classroom assignments to standardized test performance. Researchers systematically collected and analyzed classroom writing and mathematics assignments in grades 3, 6, and 8 from randomly selected and control schools over the course of three years. In addition, they evaluated student work generated by the various assignments. Finally, the researchers examined correlations among the nature of classroom assignments, the quality of student work, and scores on standardized tests. Assignments were rated according to the degree to which they required "authentic" intellectual work, which the researchers described as follows:

Authentic intellectual work involves original application of knowledge and skills, rather than just routine use of facts and procedures. It also entails disciplined inquiry into the details of a particular problem and results in a product or presentation that has meaning or value beyond success in school. We summarize these distinctive characteristics of authentic intellectual work as construction of knowledge, through the use of disciplined inquiry, to produce discourse, products, or performances that have value beyond school. (pp. 14–15)

This study concluded that

Students who received assignments requiring more challenging intellectual work also achieved greater than average gains on the Iowa Tests of Basic Skills in reading and mathematics, and demonstrated higher performance in reading, mathematics, and writing on the Illinois Goals Assessment Program. Contrary to some expectations, we found high-quality assignments in some very disadvantaged Chicago classrooms and [found] that all students in these classes benefited from exposure to such instruction. We conclude, therefore, [that] assignments calling for more authentic intellectual work actually improve student scores on conventional tests. (p. 29)[3]

Readers will immediately recognize the parallels with UbD. The instructional methods that were found to enhance student achievement are basic elements of the pedagogy in the 3-stage planning model. As in the researchers' conception of authentic intellectual work, UbD instructional approaches call for the student to construct meaning through disciplined inquiry. Assessments of understanding call for students to apply their learning in authentic contexts and explain or justify their work.

We have been asked, "Are you then saying that a more concerted effort to 'teach to the test' *lowers* scores?" No, we are not. Teaching to the test clearly has *some* effect, particularly if prior to such practice there was little attention to common standards and a focus on results. Scores do increase in the short run when a school or district focuses more carefully on a common goal. No surprise here: Greater attention to an outcome will improve performance on

any measure. But once the test particulars are figured out and students have become familiar with the test format and test-taking skills, there is rarely long-term progress. More ominously, the scores typically drop when the test is altered or re-normed.[4]

Finally, consider common sense evidence for our claim. Do we see more "teaching and assessing for understanding" in the *worse* performing schools? Do we see students more involved in slavish practicing of state and national tests in the most *high-achieving* schools? On the contrary, during the past 15 years of work with hundreds of schools and districts throughout the United States and Canada (including some of the best public and private schools in the country), we have observed more in-depth teaching and demanding assessment in the higher-performing schools. In contrast, within the lower-performing schools we found drill and practice orientations ostensibly designed to raise standardized test scores—often at the expense of more meaningful learning and lasting performance gains.

The bottom line is that we should be teaching to standards and developing the kinds of complex assessments reflected in the language of the standards, not the audit.

Misconception 2: "Yes, but . . . we have too much content to cover."

Teachers of students from kindergarten to graduate school wrestle with the reality described in the familiar phrases "information age" and "knowledge explosion." They face the challenge daily—there is simply too much information, and it is expanding too rapidly, to ever hope to "cover" it all.

In theory, the standards movement promised a solution to the problem of information overload by identifying curricular priorities. Content standards were intended to specify what is most important for students to know and be able to do, thus providing a much-needed focus and prioritization for curriculum, instruction, and assessment. In practice, content standards committees at the national, state, and district level often worked in isolation to produce overly ambitious lists of essentials for their disciplines. Rather than streamlining the curriculum, the plethora of standards in many states contributed to the overload problem.

The stress is needlessly heightened by the propensity of many teachers to treat textbooks as their teaching obligation. Those teachers have a basic misunderstanding and we can correct it: They need to use the textbook as a resource, not the syllabus. A course has certain priorities, framed as performance goals and understandings. It makes no sense to assume that everything in the textbook should be taught in class or learned by all students. U.S. textbook publishers try to cover the waterfront in order to appease 50 state textbook adoption committees, national subject-area organizations, and various special-interest groups. The result is invariably superficial treatment of the entire array of expert knowledge.

Seeing overloaded textbooks and long lists of content standards frequently leads to a fundamental misconception on the part of many teachers that their job is to cover lots of content. The perceived need to "cover" is typically based upon two implicit assumptions that we think are quite unfounded: (1) if I "teach" it (e.g., talk about it and assign some work on it), it will be adequately learned for tests; and (2) if I don't address it in a didactic way, it won't be learned.

As we have noted throughout the book, the Expert Blind Spot is hard at work here. "Teaching by mentioning" is unlikely to ensure that novices recall, much less understand, the key ideas and core processes of the subject. A superficial and disconnected teaching of information simply cannot yield optimal results on any test. We are once again confusing the teaching with the learning.

Interestingly, when teachers maintain that they are *required* to march through texts and syllabi (irrespective of the degree of student understanding or the learning results), they often cite reports of external pressures from supervisors. We have never been able to trace such reports to the administrative source nor have we found a supervisor who claimed to have issued such an edict. Our inquiries into these claims revealed that teachers were often interpreting a principal's or supervisor's focus on test scores as an *implied* request to stick closely to textbooks and test preparation as the sole strategy.

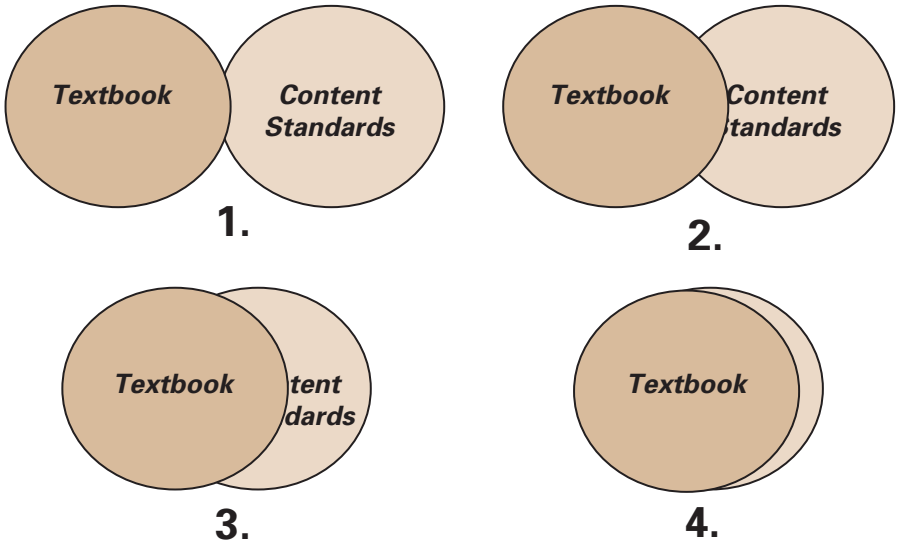
The obligation to state content standards raises an important question regarding the fit between state standards and a nationally marketed textbook or commercial resource. Ask teachers to review their textbook against state or district content standards to determine the degree of correlation. Ask them to choose the illustration in Figure 13.1 that best represents the relationship between their standards and the textbook.

In the absence of a perfect correlation between the textbook and the syllabus, the textbook should at best serve as only one of many resources, *not* the syllabus. The illustrations labeled 2 and 3 suggest that a portion of the textbook's content does not contribute to learning the standards (will not need to be learned), but that other resources will be needed.

A more disturbing exercise is to seek and find the few independent reviews of textbooks. The most thorough reviews are from Project 2061 of the American Association for the Advancement of Science (AAAS) and relate to texts for high school biology, middle school science, and algebra. The results are alarming.

> *Today's high-school biology textbooks fail to make important biology ideas comprehensible. . . . [The president of the National Academy of Sciences notes that] "sadly, it appears that our textbooks continue to be distorted by a commercial textbook market that requires that they cover the entire range of facts . . . thereby sacrificing the opportunity to treat the central concepts in enough depth to give our students a chance to truly understand them."* [5]
>
> *Not one of the widely used science textbooks for middle school was rated satisfactory. . . . "Our students are lugging home heavy texts full of disconnected*

Figure 13.1
Correlation Between Textbooks and Standards



facts that neither educate nor motivate them," said Dr. George Nelson, Director of Project 2061. . . . "This study confirms our worst fears about the materials used to educate our children in the critical middle grades."[6]

Even if good textbooks are available, it is simply a misunderstanding to claim that a teacher's job is to teach the text. The job of design, instruction, and assessment is to shape a syllabus in light of content standards, intellectual priorities, and student needs and interests to achieve explicit goals. Thus, the textbook should serve as one resource among many in the service of meeting the standards. The textbook is a reference book. Its purpose is to summarize knowledge—not unlike the encyclopedia. Treating the textbook as the syllabus ensures a lack of purpose and coherence to the overall design. Treating the textbook as the course of study is akin to marching through the encyclopedia from *A* to *Z*. Logical and efficient, yes; purposeful and effective, no.

Why isn't this misconception seen more clearly? Perhaps because school systems fail to adequately address the essential question, "What *is* my job?" when hiring, supervising, and evaluating staff. Few systems have performance-based job descriptions. Most curricula in middle and high schools assume the textbook is the syllabus. School staff members are typically hired on the basis of credentials to fill an available slot (U.S. history, 3rd grade). So, without further clarification, it is easy to fall into thinking that the textbook is the job. It is fair to say, though, that even with this lack of clarity, we've *never* seen a district contract that specifies that a teacher's job is to get through the maximum number of textbook pages. We *do* know that 49 of 50 states have established state content standards and that teachers in those states are expected to teach to them.[7]

What we learn from the international studies

The Third International Mathematics and Science Study (TIMSS), conducted in 1995, supports this view. Researchers tested mathematics and science achievement of students in 42 countries at three grade levels (4, 8, and 12). TIMSS was the largest, most comprehensive and rigorous assessment ever undertaken. Although the outcomes of TIMSS are well known—U.S. students were outperformed by students in most other industrialized countries (Martin, Mullis, Gregory, Hoyle, & Shen, 2000)—the results of the less-publicized companion TIMSS teaching study offer intriguing explanatory insights concerning the issue of content coverage in textbooks. In short, in-depth teaching for understanding using a problem-based approach supported by small texts provides far better results than the typical overloaded-textbook-focused U.S. approach.

Following an exhaustive analysis of classroom teaching in the United States, Japan, and Germany, researchers present striking evidence of the benefits of teaching for understanding in optimizing performance. For example, data from the TIMSS tests and instructional studies clearly show that, although Japanese teachers teach fewer topics in mathematics, their students achieve better results. Rather than randomly covering many discrete skills, teachers in Japan state that their primary aim is to develop conceptual understanding in their students. They emphasize depth versus superficial coverage. Although teachers in Japan cover less ground in terms of discrete topics or pages in a textbook, they emphasize problem-based learning in which rules and theorems are derived and explained by the students, thus leading to deeper understanding (Stigler & Hiebert, 1999).

Despite the fact that mathematics teachers in Japan cover fewer topics, their students achieve better results on tests. Rather than saying that their aim is the development of many discrete skills, these teachers report that their aim is conceptual understanding, and their teaching practices reflect these aims, which contrasts sharply with the views U.S. teachers have of their job. In Japan, the goal of a lesson is for students to develop mathematical thinking, whereas in other countries the goal is to acquire a specific mathematical procedure. Researchers summarized the differences between typical 8th grade mathematics lessons in Japan, Germany, and the United States:

> *The emphasis on understanding is evident in the steps typical of Japanese eighth-grade mathematics lessons:*
>
> > *Teacher poses a complex, thought-provoking problem.*
> > *Students struggle with the problem.*
> > *Various students present ideas or solutions to the class.*
> > *The teacher summarizes the class' conclusions.*
> > *Students practice similar problems.*
>
> *In contrast, the emphasis on skill acquisition is evident in the steps common to most U.S. and German mathematics lessons:*

Teacher instructs students in a concept or skill.
Teacher solves example problems with the class.
Students practice on their own while the teacher assists individual students. (U.S. Department of Education, 1999).

Teachers in Japan emphasize problem-based learning in which rules and theorems are often derived, not merely stated and reinforced through drill. Forty-two percent of their 8th grade math classes involved student presentation of possible alternative solutions to problems as opposed to only 8 percent in U.S. classrooms. Students in Japan spend 44 percent of class time trying to induce the idea to be learned from problems; students in U.S. classrooms spend less than 1 percent on that skill. In contrast, 95 percent of the time in U.S. classrooms is spent practicing a procedure to be learned, something that happens only 40 percent of the time in a Japanese classroom.

In a related finding, the researchers noted that U.S. teachers address far more topics in mathematics and science than do their international colleagues. They also make far fewer connections to other lessons—96 percent of middle school teachers in Japan made such links as compared with only 40 percent of teachers in the United States:

One way to measure coherence is to look for threats to coherence, features of lessons that make it difficult to design "design" and sustain a smoothly developing story. Threats include things like switching topics frequently, or being interrupted by outside intrusions. We found that U.S. lessons contained significantly more topics than Japanese lessons, and significantly more switches from topic to topic than did both German and Japanese lessons. (Stigler & Hiebert, 1999, p. 61)

Japanese teachers go into far greater depth than U.S. teachers:

We defined "developed" quite generously to include cases in which the concept was explained or illustrated, even with a few sentences or brief example. We found that one-fifth of the topics in U.S. lessons contained developed concepts, while four-fifths contained only stated concepts. . . . This distribution was nearly reversed in . . . Japan. (p. 60)

One of the reasons we dubbed American teaching "learning terms and practicing procedures" is that lessons in United States seemed to place greater emphasis on definitions of terms and less emphasis on underlying rationale. When we counted the number of definitions presented in all lessons, we found that there were about twice as many in the United States as in Germany or Japan. (p. 58)

Teaching versus learning

As the discussion of definitions implies, an assumption often hidden under the "need to cover" lies in thinking that everything that we want learned must be taught and that teaching the key facts is what causing learning is. This is

simply not true, as a moment's reflection on assignments that are grounded in student research, discussion, and actual performance—using facts to *do* the subject—indicate. Much of what we aim for students to learn is gained from well-designed work and as they make the effort to understand (perhaps by watching the artist, athlete, and computer scientist). Many critics of E. D. Hirsch's work misunderstand—nowhere does he advocate the direct teaching of all those core facts, only that the student learns them if he is to be equipped for the cultural literacy needed for high-level intellectual performance. (*Understanding by Design* has been successfully used in Core Literacy schools based on Hirsch's work as well as alternative project-based schools at the opposite end of the political spectrum.) Hirsch (1988) did not say that the famous list of facts is all that matters or that it must be taught didactically:

> *The extensive curriculum can be taught in a highly formal traditional school or in an informal progressive school. Any sort of school can find ways of incorporating these minimal contents in its courses. . . . The intensive curriculum, though different, is equally essential. Intensive study encourages a fully developed understanding of a subject, making one's knowledge of it integrated and coherent. . . . To understand how isolated facts fit together in some coherent way, we must always acquire mental models of how they cohere, and these schemata can only come from detailed, intensive study and experience. (pp. 128–129)*

As we noted in discussing understandings, teaching big ideas as information to be recalled must fail. Big ideas—justice, irrational numbers, irony—are inherently abstract or even counterintuitive to the naïve learner. They need uncoverage—intensive study. In fact, overly didactic teaching is a major cause, we believe, of the student misunderstanding described in earlier chapters.[8]

Then, it surely is not too controversial to say that the job of teaching is to optimize student learning of what is worthy—not to "cover" a book, nor to "teach, test and hope for the best," irrespective of results. We think that backward design, from content and performance standards (and the assessments they imply), not textbook layout, is the best way to honor that obligation.

Our own informal research findings are germane here. Recall the most common answers given by educators when they are asked to reflect on the qualities of the best instructional designs:

- Clear performance goals, based on a genuine and explicit challenge
- Hands-on approach throughout; far less front-loaded teaching than typical
- Focus on interesting and important ideas, questions, issues, problems
- Obvious real-world application, hence meaning for learners
- Powerful feedback system, with opportunities to learn from trial and error
- Personalized approach, with more than one way to do the major tasks, and room for adapting the process and goal to style, interest, need
- Clear models and modeling
- Time set aside for focused reflection

- Variety in methods, grouping, tasks
- Safe environment for taking risks
- Teacher role resembles that of a facilitator or coach
- More of an immersion experience than a typical classroom experience
- Big picture provided and clear throughout, with a transparent back-and-forth flow between the parts and the whole

The formal research on learning lends further support to the common sense of educators. In the most exhaustive summary of learning conducted in recent years, the authors of *How People Learn* make clear that more coverage does not equal more learning. Three findings form the basis of that book:

1. Students come to the classroom with preconceptions. If their initial understanding is not engaged, they may fail to grasp the new concepts and information.

2. To develop competence in an area of inquiry, students must (a) have a deep foundation of factual knowledge, (b) understand facts and ideas in the context of a conceptual framework, and (c) organize knowledge in ways that facilitate retrieval and application.

3. A metacognitive approach to instruction can help students learn to take control of their learning by defining learning goals and monitoring their progress in achieving them.

In short, "evidence from research indicates that when these three principles are incorporated into teaching, student achievement improves" (Bransford, Brown, & Cocking, 2000, p. 21).

What are some of the key implications for design and instruction? Here we'll highlight a few of the most relevant proposed by the authors, on transfer and understanding:

> *A major goal of schooling is to prepare students for flexible adaptation to new problems and settings. Students' ability to transfer what they have learned to new situations provides an important index of adaptive, flexible, learning. (p. 235)*
>
> *Knowledge of a large set of disconnected facts is not sufficient. To develop competence in an area of inquiry, students must have opportunities to learn with understanding. Deep understanding of subject matter transforms factual information into usable knowledge. . . . A key finding in the . . . literature is that organizing information into a conceptual framework allows for greater "transfer." (pp. 16, 17)*
>
> *Learning with understanding is more likely to promote transfer than simply memorizing information. . . . Many classroom activities . . . focus on facts or details rather than larger themes of causes and consequences.*
>
> *Students develop flexible understanding of when, where, why, and how to use their knowledge to solve new problems if* they learn how to extract underlying themes and principles from their learning experiences. *Understanding*

> *how and when to put knowledge to use . . . is an important characteristic of expertise. Learning in multiple contexts most likely affects this aspect of transfer. (p. 236 [emphasis added])*
>
> *Superficial coverage of all topics must be replaced with in-depth coverage of fewer topics that allows for key concepts in that discipline to be understood. (p. 20)*

Despite the typical U.S. educational mantra and fears about having to teach to the test, coverage—with equal attention to each little fact or subskill (as opposed to a focus on ideas and performance challenges that give meaning to the facts and subskills)—simply does not work to maximize test scores.

Misconception 3: "Yes, but . . . this work is too hard and I just don't have the time."

Even if we are able to convince educators that the first and second "Yes, but . . ." statements are based on misunderstandings and sustained mostly by habit, a third argument invariably arises: The time needed to do all of this work is not currently available. We agree, in part. On the surface this statement is not a misconception. Yes, aligning curriculum with state standards, identifying "big ideas," creating essential questions, designing more authentic assessments, developing plans to teach for understanding in engaging ways, analyzing the resulting student work, and conducting action research to validate interventions is very challenging work, and time must be given to it. And no, individual teachers do not have all the time needed for this difficult work (if it is to be done well). But we need to work smarter, not merely harder or more.

To work smarter we have to realize that a few other misunderstandings lurk as somewhat unconscious assumptions: (1) that *each* teacher, *each* school or *each* district must climb this mountain alone; (2) that the time required must come directly from teaching time, which (we agree) is already in short supply; (3) each standard and benchmark must be addressed separately, in dozens of units designed from scratch; and (4) that "hard and time-consuming" is a *bad* thing.

Building ongoing collaborative research and development into the job

As the exercise on best design suggests, deeper teacher understandings can often be best developed through local study groups and action research. We must apply what the list says and what the book proposes to teacher understanding: An *in-depth* investigation of big ideas in learning is what matters, and school must make that learning more central to professional development and the job description. At the heart of so many of these problems is a variant of the teacher blind spot: "I taught it, so they must have learned it; if I teach more stuff, they will learn more." No. Left to ourselves, the habit of

coverage is always likely to seem more defensible than it is. We have to better *understand* learning. We must develop "perspective" and learn to better "apply" sound theoretical "explanations" and "interpretations" of educational research to our work.

Teachers, teams, departments, and entire faculties must ask themselves *each* year: What approaches to curriculum design, teaching, and assessing actually yield the greatest student learning, regardless of our habits and attitudes? We learn from the answers to those questions that we must practice in professional development what Understanding by Design says about understanding: We must uncover, not cover the big ideas, through ongoing inquiry and discussion.

But mindful of both real issues of time and the need for intensive study, let the local research be small in scope but deep—focused on a single unit a year. One unit, designed in collaboration with others, tried out and adjusted a few times per year, with intensive analysis of the student work—surely such a process is possible within existing time allotted for in-service days and team meetings. Consider an analogy to see both the feasibility and value of proceeding this way: How many busy educators prepare gourmet meals *every evening* during the school year? We chuckle at the thought. Even the avid cooks among us don't have the time or energy. But a few times per year, maybe more, we do engage in more elaborate home dining (e.g., a family holiday dinner) that requires more extensive planning, preparation time, and attention to presentation than do typical daily meals. Let it be a job requirement of teaching, backed by supervision, that one "gourmet" unit per year must be developed, collected, reviewed and shared. (Imagine, then, the resulting school or district curriculum "cookbook" ten years hence!)

Such an incremental approach, grounded in the development of design exemplars from which we can all learn, was the central recommendation in *The Teaching Gap.* In addition to the instructional differences between teachers in Japan and the United States, the researchers noted another important difference in the ongoing teacher education in the two countries. In Japan, teacher education seeks depth, not breadth, uncoverage not coverage, learning on the job, not "teaching" of new techniques. For decades, teachers in Japan have used a process known as lesson study, whereby they regularly work in small teams to develop, teach, and refine one research lesson per year. They share the results of their action research and concomitant lesson designs not only with their colleagues in staff meetings, but in regional lesson fairs so that other teachers will benefit from their insights.

We stress that reforming professional development for teachers is the only guaranteed way to improve standard practice and professionalism among all teachers:

> *Another important benefit of the collaborative nature of lesson study is that it provides a benchmarking process that teachers can use to gauge their own skills. . . . At the same time, the collaborative nature of lesson study balances the self-critiquing . . . with the idea that improved teaching is a joint*

process. . . . Problems that emerge are generally attributed to the lesson as designed by the group. . . . It thus becomes possible for teachers to be critical. (Stigler & Hiebert, 1999, p. 125)[9]

This process of collaborative unit and lesson design, refinement, and regional sharing is reflected in the UbD peer-review process. Specific information, directions, and samples are in the *Understanding by Design Professional Development Workbook* (McTighe & Wiggins, 2004).

What is odd about the lack-of-time complaint is that it is only partly true. Every school system devotes at least 12 hours a year to professional development days and approximately 16 hours to staff, grade-level, and departmental meetings. Just imagine what could be accomplished by rethinking those hours and devoting half of them to some form of required lesson study as a job requirement, embedded in grade-level and departmental meeting schedules, as well as the in-service days. Over time, action research would become part of the obligation of all teams and departments, with annual reports issued as to achievement targets tackled, research and development undertaken, results found, and new inquiry proposed for the future.

Consider the following example of how this approach works. Imagine that teachers in your school or district had the opportunity once every three years to take part in a regional summer curriculum design workshop. They would be invited to bring the best (e.g., most engaging and effective) unit that they teach (connected, of course, to state or district content standards). They would join with one or two other teachers of the same subject and level who have identified a similar unit topic, and work with the guidance of a content expert to prepare a "gourmet" unit. Their work in progress would be reviewed against a set of curriculum design standards (such as those in *Understanding by Design*), and they would make adjustments based on the feedback received from peers and experts. They would then enter their best ideas on a computer in an agreed-upon format such as the UbD 3-stage backward design template, and as happens with the UbD Exchange (http://ubdexchange.org).

During the following school year, they would field-test their enhanced unit and collect student work as evidence of the results. They would meet during the year (perhaps during a scheduled in-service day) to collectively evaluate the student work, and make needed adjustments to their unit design. Their completed design would be eligible for regional review by content experts (based on the design standards and the results from student work). Those units that were deemed exemplary would be made available to other educators through the electronic database. We have helped many school faculties develop such a system over the past five years.[10]

A misconception in the way of more such collaboration (fostered in part by local culture, where teachers are dysfunctionally isolated from one another) is that we often presume that content standards and benchmarks need to be addressed discretely, one at a time through narrow targeted lessons, by each teacher, isolated in a classroom. That understandably breeds the feeling that

the work is too much for any one of us to handle, but the premise is flawed. This confusion relates back to the first "Yes, but . . ." argument and the problem of the face validity of the state tests. Standardized tests typically sample the standards one at a time through decontexualized (aptly named) "items." Thus, the look and feel of the tests and the lists of standards often misleadingly suggest that we should teach to the standards one bit at a time, as if each standard, benchmark, and indicator is of equal importance.

On the contrary, we are back to the beginning of Understanding by Design: the 3-circle graphic whereby we set priorities around big ideas and core tasks, derived from the standards. Then, when units involve rich and in-depth work, culminating in complex performance, dozens of standards are addressed simultaneously, in appropriate hierarchical order—and with more coherence from the learner's perspective. The challenge at the local level is not to design a lesson per indicator, but to design rich units that ultimately address all standards and clearly signal to students the priorities. This is a problem that is solvable by better unpacking of standards, curriculum writing, mapping, and data collection.

We contend that all such action research will yield four distinct benefits:

1. Walking the Talk. By applying standards to *our own* professional work, the quality of curriculum and assessment designs is enhanced. Instead of assuming that our designs are sound because we worked hard or included activities that students enjoy, the designs must be validated against design standards. Curriculum designs that meet the standards and result in student learning are designated as exemplars, and thus, establish high standards for future curriculum work.

2. Mental Templating. The logic of backward design calls for clarity about desired results and needed assessment evidence *before* identifying learning activities or selecting resources. When teachers use a backward design template to design curriculum units, they develop a productive mental model for planning that helps to avoid the twin problems of activity-oriented and coverage-oriented curricula. Such a design process is particularly valuable for new teachers who have not yet developed a file cabinet of favorite activities or been fully seduced by reliance on textbooks.

3. Working Smarter Using Technology. Most educators are obligated to teach to their state content standards, so why shouldn't statewide sharing be the norm? Because state standards in the various disciplines are more similar than different, couldn't this sharing be extended to a national level? We believe so. Rather than each teacher, school, and district unnecessarily reinventing the same wheels, this approach provides a mechanism for working smarter using a searchable database of validated units. We need not feel guilty that we do not prepare gourmet meals every day. Focusing regularly on exemplars (be they recipes or curriculum designs) is good for everyone. We can thus devote our energies to developing one or two high-quality units, and develop increasingly

higher standards and more refined design skills as we work. And, as with cookbooks, it is far smarter to share our designs so that everyone can benefit from proven recipes.

4. Enhancing Professional Conversations. In addition to higher-quality curriculum products, the *process* of shared design work provides rich professional development. Responses from teachers working in cross-district design teams (as part of regional and state consortia) have confirmed the value of the experience. Unlike one-size-fits-all staff development sessions on generic topics, this design work concentrates on the unique aspects of teaching and assessing specific content topics and results in tangible products of immediate value to teachers. Conversations focus on matters at the heart of the profession: What are the big ideas that we want students to understand? How will we know that they really learned this? What does it mean to meet these standards? What teaching and learning experiences are most engaging and effective? What does student work reveal about the strengths and needs of our curriculum and instruction? Given the limited available time for professional development, it is imperative that it be results-oriented in such a manner—as opposed to coverage of educational trends by outside speakers.

Time-consuming hard work is not a bad thing. It is a good and vital thing, as the quote from Plato's *Republic* at the opening of the chapter suggests. Learning, true learning, is always difficult. It *always* upends old learning, leading to disequilibrium and resistance. We have found that many educators have a paradoxical resistance to learning—especially teachers used to working alone and thinking that smooth control of all that happens in their space, based on their habits, is what matters most. Perhaps the best reason to redesign schools around ongoing collaborative research is that this is the only way to overcome teacher resistance to changing habits, a timidity about experimentation, and a fear of criticism and failure. There is greater courage—and helpful peer pressure to *learn*—when a group of teachers works together to do research into their individual and collective practice.

The six facets are involved here. Teachers need work that will develop greater empathy and self-understanding if they are to truly understand how to cause learning. The blind spot hides from us the pain of all learning, not just the likelihood of student misconception or the individual differences and needs of learners. We are not teachers; we are causes of and students of learning. The job should therefore require that we get and remain "inside" how learning works, to constantly remind ourselves of how difficult it really is. School should require teachers to do action research so that they constantly *feel* what it is like to learn, to be reminded that real learning is always frightening, frustrating, and able to cause self-doubt, regardless of age or talent. If the job and schedule make us think of ourselves as only teachers instead of also as model learners, we miss vital opportunities to make education more honest, invigorating, and self-correcting for everyone, adult and child. The time needed for this work should not be construed as extra but as essential.

Conclusion

We have considered a few widely held views about the obstacles to designing, teaching, and assessing for understanding in a world of external accountability, and we have attempted to reveal their underlying misconceptions. We have suggested that ideas from Understanding by Design are central to improving performance on external measures while preserving intellectual engagement for staff and students: (1) teach and assess for understanding of big ideas and mastery of core subject-area tasks; (2) apply design standards to review and refine local curriculum and assessment, as part of on-going local research and development embedded in the job.

We do not ask or expect you to take us at our word. We know from experience that habits and misunderstandings are rarely overcome by argument—by our covering them! No, the claims in this chapter and in the book as a whole need to be locally uncovered and discussed, tested, argued, and explored by you in your own setting if they are to be accepted (or rejected) on rational grounds. This, too, was, a key conclusion about U.S. school reform drawn by the authors of *The Teaching Gap*:

> *Because teaching is complex, improvements in teaching will be most successful if they are developed in the classrooms where teachers teach and students learn. . . . What works in one classroom might or might not work in another classroom. Ideas for improvement that come from afar—including, for example, what we've learned from Japanese lessons—will need to be tested and adapted to our own local classrooms. (Stigler, & Hiebert, 1999, p. 134)*

We challenge you to investigate what understanding is and isn't, how to best teach for it, and how to best assess for it—all this, in your world of particular standards, tests, and students. All the research in the world means little if you cannot see it at work in *your* classes, with *your* students. Understanding this book means doing the work of trying out the ideas in the book. That's what lesson study sets in motion.

It is our hope that by uncovering some of these often-heard pessimistic claims, we may encourage a more proactive stance by school faculties and district leaders toward what you *can* do to improve learning, regardless of the setting in which you find yourself and the hard work required. The research findings are heartening. Regardless of all the things about students, schools, and society that we cannot control, the things that are in our control—design, instruction, giving feedback—can still significantly affect achievement.