Key terms

A few words about terminology are in order. We talk a good deal in the book about *big ideas* that should be the focus of education for understanding. A big idea is a concept, theme, or issue that gives meaning and connection to discrete facts and skills. Here are some examples: adaptation; how form and function are related in systems; the distributive property in mathematics (whereby we can use any number of groupings and subgroupings to yield the "same" numbers); problem solving as the finding of useful models; the challenge of defining *justice;* and the need to focus on audience and purpose as a writer or speaker. In an education for understanding, a vital challenge is to highlight the big ideas, show how they prioritize the learning, and help students understand their value for making sense of all the "stuff" of content.

Educators involved in reform know that the words *curriculum* and *assessment* have almost as many meanings as there are people using the terms. In this book, *curriculum* refers to the specific blueprint for learning that is derived from *desired results*—that is, content and performance standards (be they state-determined or locally developed). Curriculum takes content (from external standards and local goals) and shapes it into a plan for how to conduct effective and engaging teaching and learning. It is thus more than a list of topics and lists of key facts and skills (the "inputs"). It is a map for how to achieve the "outputs" of desired student performance, in which appropriate learning activities and assessments are suggested to make it more likely that students achieve the desired results.

The etymology of the word suggests this: *Curriculum* is the particular "course to be run," given a desired end point. A curriculum is more than a traditional program guide, therefore; beyond mapping out the topics and materials, it specifies the most appropriate experiences, assignments, and assessments that might be used for achieving goals. The best curricula (and syllabi), in other words, are written from the point of view of the desired learnings, not merely what will be covered. They specify what the learner should have achieved upon leaving, what the learner needs to do to achieve, and what the teacher needs to do to achieve the results sought. In sum, they specify the desired output and means of achieving it, not just a list of content and activities.

By *assessment* we mean the act of determining the extent to which the desired results are on the way to being achieved and to what extent they have been achieved. Assessment is the umbrella term for the deliberate use of *many* methods of gathering evidence of meeting desired results, whether those results are state content standards or local curricular objectives. The collected evidence we seek may well include observations and dialogues, traditional quizzes and tests, performance tasks and projects, as well as students' self-assessments gathered over time. *Assessment* is thus a more learning-focused term than *evaluation,* and the two should not be viewed as synonymous. Assessment is the giving and using of feedback against standards to enable improvement and the meeting of goals. Evaluation, by contrast, is more summative and credential-related. In other words, we need not give a grade—an evaluation—to everything we give feedback to. In fact, a central premise of our argument is that understanding can be developed and evoked only through multiple methods of ongoing assessment, with far greater attention paid to formative (and performance) assessment than is typical.

By *desired results* we mean what has often been termed *intended outcomes, achievement targets,* or *performance standards.* All four terms are meant to shift our focus away from the inputs to the output: what the student should be able to know, do, and understand upon leaving, expressed in performance and product terms. *Desired result* reminds us also that, as "coaches," we will likely have to adjust *our* design and performance en route, if feedback shows that we are in danger of not achieving the successes sought.

The word *understanding* turns out to be a complex and confusing target despite the fact that we aim for it all the time. The word naturally deserves clarification and elaboration, which is the challenge for the rest of the book. For now, though, consider our initial working definition of the term: To *understand* is to make connections and bind together our knowledge into something that makes sense of things (whereas without understanding we might see only unclear, isolated, or unhelpful facts). But the word also implies doing, not just a mental act: A performance ability lies at the heart of understanding, as Bloom (1956) noted in his Taxonomy in discussing application and synthesis. To understand is to be able to wisely and effectively *use*—transfer—what we know, in context; to *apply* knowledge and skill effectively, in realistic tasks and settings. To have understood means that we show evidence of being able to transfer what we know. When we understand, we have a fluent and fluid grasp, not a rigid, formulaic grasp based only on recall and "plugging in."

When we speak of the product of this achievement—*an* understanding, as a noun—we are describing particular (often hard-won) insights. For example, we talk about scientists' current understanding that the universe is expanding or the postmodern understanding of authors as not being privileged commentators on the meaning of their books. The great challenge in teaching is to enable such subtle adult understandings to become student understandings—without reducing the understanding to a mere simplistic statement for recall. If the student gains a genuine understanding, we typically say they "*really* get it." With our help as designers and coaches, they "come to an understanding."

Yet, for years, curriculum guides have argued against framing objectives in terms of understandings. Bloom (1956) argued that the word is too ambiguous to use as a foundation for teaching goals and their assessments; hence, the writing of the Taxonomy. But an important conceptual distinction remains and needs pondering: the difference between *knowing* and *understanding.* Pinning this distinction down in theory and in practice has not been easy. We propose in the book that insufficient attention has been paid to the fact that there are *different kinds* of understandings, that knowledge and skill *do not* automatically lead to understanding, that student *misunderstanding* is a far bigger problem than we may realize, and that assessment of understanding therefore requires evidence that *cannot* be gained from traditional fact-focused testing alone.

Chapter 1

Backward Design

Design, v.,—To have purposes and intentions; to plan and execute
—*Oxford English Dictionary*

The complexity of design work is often underestimated. Many people believe they know a good deal about design. What they do not realize is how much more they need to know to do design well, with distinction, refinement, and grace.
—John McClean, "20 Considerations That Help a Project Run Smoothly," 2003

Teachers are designers. An essential act of our profession is the crafting of curriculum and learning experiences to meet specified purposes. We are also designers of assessments to diagnose student needs to guide our teaching and to enable us, our students, and others (parents and administrators) to determine whether we have achieved our goals.

Like people in other design professions, such as architecture, engineering, or graphic arts, designers in education must be mindful of their audiences. Professionals in these fields are strongly client-centered. The effectiveness of their designs corresponds to whether they have accomplished explicit goals for specific end-users. Clearly, students are our primary clients, given that the effectiveness of curriculum, assessment, and instructional designs is ultimately determined by their achievement of desired learnings. We can think of our designs, then, as software. Our courseware is designed to make learning more effective, just as computer software is intended to make its users more productive.

As in all the design professions, standards inform and shape our work. The software developer works to maximize user-friendliness and to reduce bugs that impede results. The architect is guided by building codes, customer budget, and neighborhood aesthetics. The teacher as designer is similarly constrained. We are not free to teach any topic we choose by any means. Rather, we are guided by national, state, district, or institutional standards that specify what students should know and be able to do. These standards provide a useful framework to help us identify teaching and learning priorities and guide our design of curriculum and assessments. In addition to external standards, we must also factor in the needs of our many and varied students when designing learning experiences. For example, diverse student interests, developmental levels, large classes, and previous achievements must always shape our thinking about the learning activities, assignments, and assessments.

Yet, as the old adage reminds us, in the best designs form follows function. In other words, all the methods and materials we use are shaped by a clear conception of the vision of desired results. That means that we must be able to state with clarity what the student should understand and be able to do as a result of any plan and irrespective of any constraints we face.

You probably know the saying, "If you don't know exactly where you are headed, then any road will get you there." Alas, the point is a serious one in education. We are quick to say what things *we* like to teach, what activities *we* will do, and what kinds of resources *we* will use; but without clarifying the desired results of our teaching, how will we ever know whether our designs are appropriate or arbitrary? How will we distinguish merely interesting learning from *effective* learning? More pointedly, how will we ever meet content standards or arrive at hard-won student understandings unless we think through what those goals imply for the learner's activities and achievements?

Good design, then, is not so much about gaining a few new technical skills as it is about learning to be more thoughtful and specific about our purposes and what they imply.

Why "backward" is best

How do these general design considerations apply to curriculum planning? Deliberate and focused instructional design requires us as teachers and curriculum writers to make an important shift in our thinking about the nature of our job. The shift involves thinking a great deal, first, about the specific learnings sought, and the evidence of such learnings, before thinking about what we, as the teacher, will do or provide in teaching and learning activities. Though considerations about what to teach and how to teach it may dominate our thinking as a matter of habit, the challenge is to focus first on the desired learnings from which appropriate teaching will logically follow.

Our lessons, units, and courses should be logically inferred from the results sought, not derived from the methods, books, and activities with which we are most comfortable. Curriculum should lay out the most effective ways of achieving specific results. It is analogous to travel planning. Our frameworks should provide a set of itineraries deliberately designed to meet cultural goals rather than a purposeless tour of all the major sites in a foreign country. In short, the best designs derive backward from the learnings sought.

The appropriateness of this approach becomes clearer when we consider the educational purpose that is the focus of this book: understanding. We cannot say *how* to teach for understanding or *which* material and activities to use until we are quite clear about which specific understandings we are after and what such understandings look like in practice. We can best decide, as guides, what "sites" to have our student "tourists" visit and what specific "culture" they should experience in their brief time there only if we are clear about the particular understandings about the culture we want them to take home. Only by having specified the desired results can we focus on the content, methods, and activities most likely to achieve those results.

But many teachers begin with and remain focused on textbooks, favored lessons, and time-honored activities—the inputs—rather than deriving those means from what is implied in the desired results—the output. To put it in an odd way, too many teachers focus on the *teaching* and not the *learning*. They spend most of their time thinking, first, about what they will do, what materials they will use, and what they will ask students to do rather than first considering what the learner will need in order to accomplish the learning goals.

Consider a typical episode of what might be called *content*-focused design instead of *results*-focused design. The teacher might base a lesson on a particular topic (e.g., racial prejudice), select a resource (e.g., *To Kill a Mockingbird*), choose specific instructional methods based on the resource and topic (e.g., Socratic seminar to discuss the book and cooperative groups to analyze stereotypical images in films and on television), and hope thereby to cause learning (and meet a few English/language arts standards). Finally, the teacher might think up a few essay questions and quizzes for assessing student understanding of the book.

This approach is so common that we may well be tempted to reply, What could be wrong with such an approach? The short answer lies in the basic questions of purpose: Why are we asking students to read this particular novel—in other words, what *learnings* will we seek from their having read it? Do the students grasp why and how the purpose should influence their studying? What should students be expected to understand and do upon reading the book, related to our goals beyond the book? Unless we begin our design work with a clear insight into larger purposes—whereby the book is properly thought of as a means to an educational end, not an end unto itself—it is unlikely that all students will *understand* the book (and their performance obligations). Without being self-conscious of the specific understandings about prejudice we seek, and how reading and discussing the book will help develop such insights, the goal is far too vague: The approach is more "by hope" than "by design." Such an approach ends up unwittingly being one that could be described like this: Throw some content and activities against the wall and hope some of it sticks.

Design Tip

Consider these questions that arise in the minds of all readers, the answers to which will frame the priorities of coached learning: How should I read the book? What am I looking for? What will we discuss? How should I prepare for those discussions? How do I know if my reading and discussions are effective? Toward what performance goals do this reading and these discussions head, so that I might focus and prioritize my studies and note taking? What big ideas, linked to other readings, are in play here? These are the students' proper questions about the learning, not the teaching, and any good educational design answers them from the start and throughout a course of study with the use of tools and strategies such as graphic organizers and written guidelines.

Answering the "why?" and "so what?" questions that older students always ask (or want to), and doing so in concrete terms as the focus of curriculum planning, is thus the essence of understanding by design. What is difficult for many teachers to see (but easier for students to feel!) is that, without such explicit and transparent priorities, many students find day-to-day work confusing and frustrating.

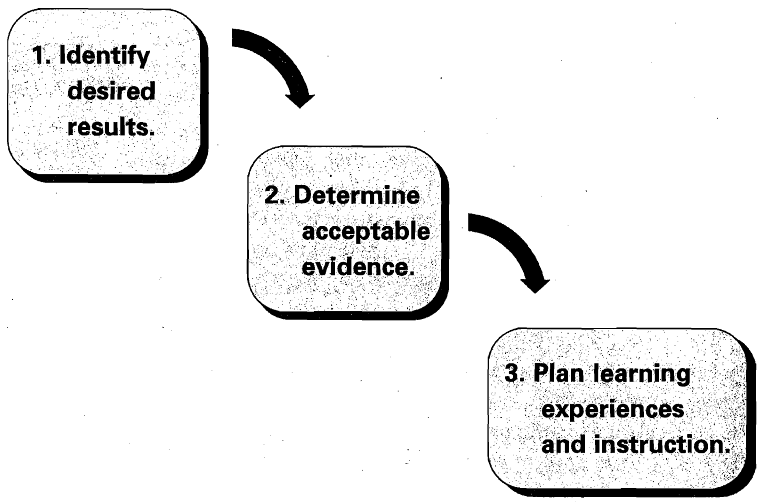
The three stages of backward design

We call this three-stage approach to planning "backward design." Figure 1.1 depicts the three stages in the simplest terms.

Stage 1: Identify desired results

What should students know, understand, and be able to do? What content is worthy of understanding? What *enduring* understandings are desired?

Figure 1.1
UbD: Stages of Backward Design



In Stage 1 we consider our goals, examine established content standards (national, state, district), and review curriculum expectations. Because typically we have more content than we can reasonably address within the available time, we must make choices. This first stage in the design process calls for clarity about priorities.

Stage 2: Determine acceptable evidence

How will we know if students have achieved the desired results? What will we accept as evidence of student understanding and proficiency? The backward design orientation suggests that we think about a unit or course in terms of the collected assessment evidence needed to document and validate that the desired learning has been achieved, not simply as content to be covered or as a series of learning activities. This approach encourages teachers and curriculum planners to first "think like an assessor" before designing specific units and lessons, and thus to consider up front how they will determine if students have attained the desired understandings.

Stage 3: Plan learning experiences and instruction

With clearly identified results and appropriate evidence of understanding in mind, it is now the time to fully think through the most appropriate instructional activities. Several key questions must be considered at this stage of backward design: What enabling knowledge (facts, concepts, principles) and skills (processes, procedures, strategies) will students need in order to perform effectively and achieve desired results? What activities will equip students with the needed knowledge and skills? What will need to be taught and coached, and how should it best be taught, in light of performance goals? What materials and resources are best suited to accomplish these goals?

Note that the specifics of instructional planning—choices about teaching methods, sequence of lessons, and resource materials—can be successfully completed only after we identify desired results and assessments and consider what they imply. Teaching is a means to an end. Having a clear goal helps to focus our planning and guide purposeful action toward the intended results.

Backward design may be thought of, in other words, as purposeful task analysis: Given a worthy task to be accomplished, how do we best get everyone equipped? Or we might think of it as building a wise itinerary, using a map: Given a destination, what's the most effective and efficient route? Or we might think of it as planning for coaching, as suggested earlier: What must learners master if they are to effectively perform? What will count as evidence *on the field,* not merely in drills, that they really get it and are ready to *perform with understanding, knowledge, and skill* on their own? How will the learning be designed so that learners' capacities are developed through use and feedback?

■ MISCONCEPTION ALERT!

When we speak of evidence of desired results, we are referring to evidence gathered through a variety of formal and informal assessments during a unit of study or a course. We are not alluding only to end-of-teaching tests or culminating tasks. Rather, the collected evidence we seek may well include traditional quizzes and tests, performance tasks and projects, observations and dialogues, as well as students' self-assessments gathered over time.

This is all quite logical when you come to understand it, but "backward" from the perspective of much habit and tradition in our field. A major change from common practice occurs as designers must begin to think about assessment *before* deciding what and how they will teach. Rather than creating assessments near the conclusion of a unit of study (or relying on the tests provided by textbook publishers, which may not completely or appropriately assess our standards and goals), backward design calls for us to make our goals or standards specific and concrete, in terms of assessment evidence, as we begin to plan a unit or course.

The logic of backward design applies regardless of the learning goals. For example, when starting from a state content standard, curriculum designers need to determine the appropriate assessment evidence stated or implied in the standard. Likewise, a staff developer should determine what evidence will indicate that the adults have learned the intended knowledge or skill before planning the various workshop activities.

Figure 1.2
1-Page Template with Design Questions for Teachers

Stage 1—Desired Results

Established Goals: G
- What relevant goals (e.g., content standards, course or program objectives, learning outcomes) will this design address?

Understandings: U
Students will understand that . . .
- What are the big ideas?
- What specific understandings about them are desired?
- What misunderstandings are predictable?

Essential Questions: Q
- What provocative questions will foster inquiry, understanding, and transfer of learning?

Students will know . . . K
- What key knowledge and skills will students acquire as a result of this unit?
- What should they eventually be able to do as a result of such knowledge and skills?

Students will be able to . . . S

Stage 2—Assessment Evidence

Performance Tasks: T
- Through what authentic performance tasks will students demonstrate the desired understandings?
- By what criteria will performances of understanding be judged?

Other Evidence: OE
- Through what other evidence (e.g., quizzes, tests, academic prompts, observations, homework, journals) will students demonstrate achievement of the desired results?
- How will students reflect upon and self-assess their learning?

Stage 3—Learning Plan

Learning Activities: L

What learning experiences and instruction will enable students to achieve the desired results? How will the design

W = Help the students know **W**here the unit is going and **W**hat is expected? Help the teacher know **W**here the students are coming from (prior knowledge, interests)?
H = **H**ook all students and **H**old their interest?
E = **E**quip students, help them **E**xperience the key ideas and **E**xplore the issues?
R = Provide opportunities to **R**ethink and **R**evise their understandings and work?
E = Allow students to **E**valuate their work and its implications?
T = Be **T**ailored (personalized) to the different needs, interests, and abilities of learners?
O = Be **O**rganized to maximize initial and sustained engagement as well as effective learning?

Figure 2.1
Knowledge Versus Understanding

Knowledge	Understanding
• The facts	• The meaning of the facts
• A body of coherent facts	• The "theory" that provides coherence and meaning to those facts
• Verifiable claims	• Fallible, in-process theories
• Right or wrong	• A matter of degree or sophistication
• I know something to be true	• I understand why it is, what makes it knowledge
• I respond on cue with what I know	• I judge when to and when not to use what I know

Understanding as transferability

> It would be impossible to over-estimate the educational importance of arriving at conceptions: that is, meanings that are general because applicable in a great variety of different instances in spite of their difference. . . . They are known points of reference by which we get our bearings when we are plunged into the strange and unknown. . . . Without this conceptualizing, nothing is gained that can be carried over to the better understanding of new experiences.
>
> —John Dewey, *How We Think,* 1933, p. 153

> Baking without an understanding of the ingredients and how they work is like baking blindfold[ed] . . . sometimes everything works. But when it doesn't you have to guess at how to change it. . . . It is this understanding which enables me to both creative and successful.
>
> —Rose Levy Berenbaum, *The Cake Bible,* 1988, p. 469

To know *which* fact to use *when* requires more than another fact. It requires understanding—insight into essentials, purpose, audience, strategy, and tactics. Drill and direct instruction can develop discrete skills and facts into automaticity (knowing "by heart"), but they cannot make us truly able.

Understanding is about *transfer,* in other words. To be truly able requires the ability to transfer what we have learned to new and sometimes confusing settings. The ability to transfer our knowledge and skill effectively involves the capacity to take what we know and use it creatively, flexibly, fluently, in different settings or problems, on our own. Transferability is not mere plugging in of previously learned knowledge and skill. In Bruner's famous phrase, understanding is about "going beyond the information given"; we can create new knowledge and arrive at further understandings if we have learned with understanding some key ideas and strategies.

What is transfer, and why does it matter? We are expected to take what we learned in one lesson and be able to apply it to other, related but different situations. Developing the ability to transfer one's learning is key to a good education (see Bransford, Brown, & Cocking, 2000, pp. 51ff). It is an essential ability because teachers can only help students learn a relatively small number of ideas, examples, facts, and skills in the entire field of study; so we need to help them transfer their inherently limited learning to many other settings, issues, and problems.

Consider a simple example from sports. When we grasp the idea that on defense we need to close up available space for the offense, we can use that understanding to adapt to almost *any* move members of the other team make, not just be limited to the one or two positionings we were taught in a three-on-three drill.

And because the big idea of "constraining offensive space" *transfers* across sports, it is equally applicable in soccer, basketball, hockey, water polo, football, and lacrosse. The same is true in math or reading: To get beyond mere rote learning and recall, we have to be taught and be assessed on an ability to see patterns, so that we come to see many "new" problems we encounter as variants of problems and techniques we are familiar with. That requires an education in how to problem solve using big ideas and transferable strategies, not merely how to plug in specific facts or formulas.

Big ideas are essential because they provide the basis for the transfer. You must learn that a single strategy underlies all possible combinations of specific moves and settings, for example. The strategy is to get someone on your team open, using various moves and fakes—regardless of what the other team does or whether it looks exactly like what you did in practice. In academics, you must learn to transfer intellectual knowledge and skill:

> *Transfer is affected by the degree to which people learn with understanding rather than merely memorize sets of facts or follow a fixed set of procedures. . . . Attempts to cover too many topics too quickly may hinder learning and subsequent transfer. (Bransford, Brown, & Cocking, 2000, pp. 55, 58)*

This is an old idea, famously framed by Whitehead (1929) almost 100 years ago in his complaint about "inert ideas" in education:

> *In training a child to activity of thought, above all things we must beware of what I will call "inert ideas"—that is to say, ideas that are merely received into the mind without being utilized or tested, or thrown into fresh combinations. . . . Education with inert ideas is not only useless: it is above all things, harmful. . . . Let the main ideas which are introduced be few and important, and let them be thrown into every combination possible. (pp. 1–2)*

Transfer is the essence of what Bloom and his colleagues meant by application. The challenge is not to "plug in" what was learned, from memory, but modify, adjust, and adapt an (inherently general) idea to the particulars of a situation:

> *Students should not be able to solve the new problems and situations merely by remembering the solution to or the precise method of solving a similar problem in class. It is not a new problem or situation if it is exactly like the others solved in class except that new quantities or symbols are used. . . . It is a new problem or situation if the student has not been given instruction or help on a given problem and must do some of the following. . . . 1. The statement of the problem must be modified in some way before it can be attacked. . . . 2. The statement of the problem must be put in the form of some model before the student can bring the generalizations previously learned to bear on it. . . . 3. The statement of the problem requires the student to search through memory for relevant generalizations. (Bloom, Madaus, & Hastings, 1981, p. 233)*

Knowledge and skill, then, are necessary elements of understanding, but not sufficient in themselves. Understanding requires more: the ability to thoughtfully and actively "do" the work with discernment, as well as the ability to self-assess, justify, and critique such "doings." Transfer involves figuring out which knowledge and skill matters *here* and often adapting what we know to address the challenge at hand.

All of the cases we've discussed here illustrate the importance of confronting students with a real problem for thought if understanding is to be called for and awakened. This is very different from giving students lessons and tests that merely require taking in and recalling from memory, based on highly cued exercises in which learners simply plug in what is unambiguously required. (See Chapters 6 through 8 for further discussions on crafting understandings and meaningful assessments.)

The failure of even our best students to transfer their learning is evident in many areas but is most striking in mathematics. Consider the following examples of test items, all of which are testing the same idea (in each case, approximately *two-thirds* of the tested students did not correctly answer the question):

From the NAEP 12th grade mathematics test:
What is the distance between the points (2,10) and (-4, 2) in the xy plane?

- ☐ 6
- ☐ 8
- ☐ 10
- ☐ 14
- ☐ 18

From a *Boston Globe* article on the Massachusetts MCAS 10th grade math scores:
The hardest question on the math section, which just 33 percent got right, asked students to calculate the distance between two points. It was a cinch—if students knew that they could plot the points and use the Pythagorean theorem, a well-known formula to calculate the hypotenuse of a right triangle if the lengths of two legs are given. The sixth-hardest math question, which only 41 percent of students got right, also required use of the Pythagorean theorem. "It seems applying the Pythagorean theorem was a weakness for kids," said William Kendall, director of math for the Braintree public schools. "These weren't straightforward Pythagorean theorem questions. They had to do a little bit more." (Vaishnav, 2003)

All three problems require students to transfer their understanding of the Pythagorean theorem to a new situation. It is likely that most students in the United States could not do it, despite the fact that *every* set of state standards identifies a grasp of the Pythagorean theorem as a key desired result.

We can apply *our* understanding to this news without too much difficulty, based on what has been said thus far. We surmise that the $A^2 + B^2 = C^2$ theorem is taught as a fact, a rule for making certain calculations when confronted with a known right triangle and simple tasks. Remove a few blatant cues, however, and students cannot transfer their learning to perform with understanding. Is it any wonder, then, that students do not *understand* what they supposedly *know?* And what few educators seem to realize, therefore, is that drilling students for state tests is a *failing* strategy.

The Expert Blind Spot

> Teaching specific topics or skills without making clear their context in the broader fundamental structure of a field of knowledge is uneconomical.
> —Jerome Bruner, *The Process of Education,* 1960, p. 31

Understanding the importance of transfer can help us make sense, then, of those educators, like Bruner, who claim that typical coverage is "uneconomical." How can he say this? It seems so manifestly false: Teaching for understanding is perhaps more *effective,* but how can it possibly be more *efficient?*

Can't we address far more content through didactic teaching and textbook coverage than we can by setting up inquiry-based work to help students come to deeper understanding of the material on their own?

But this confuses the *teaching* with the *learning*. Consider Bruner's three reasons for why a traditional coverage approach is uneconomical in the long run:

> *Such teaching makes it exceedingly difficult for the student to generalize from what he has learned to what he will encounter later. In the second place, [such] learning . . . has little reward in terms of intellectual excitement. . . . Third, knowledge one has acquired without sufficient structure to tie it together is knowledge that is likely to be forgotten. An unconnected set of facts has a pitiably short half-life in memory. (Bruner, 1960, p. 31)*

In other words, we as educators fail to understand understanding when we think that coverage works. What we call the Expert Blind Spot is hard at work, causing us to confuse what we (or textbook authors) talk about with the active meaning-making required by the learner to grasp and use meaning. This habitual response by so many of us amounts to saying, "If I cover it clearly, they will 'get it' and be able to call upon it in the future. The more I cover, therefore, the more they will learn, and the better they'll do on the tests."

What we hope you see by the book's end, however, is that this widely held assumption is false; the "yield" from coverage is quite low for most students:

> *More than 30 years ago, medical educators conducted a study on what first-year medical students remembered of the thousands of new terms that they'd memorized in their first-year gross anatomy course. They were tested and retested over time.* The curve that matched most closely to their forgetting of gross anatomy was the same shape as discovered in Ebbinghaus's classic study of memory for nonsense syllables a century ago. *The publication of data like these made a mark in the world of medical education. The teaching of anatomy has since changed radically in schools of medicine. (Shulman, 1999, p. 13 [emphasis added])*

Coverage leaves students with no sense of the whole that seems so obvious to the expert—all but the few most able students will get lost, and perhaps alienated.

Teachers do not optimize performance, even on external tests, by covering everything superficially. Students end up forgetting or misunderstanding *far more than is necessary,* so that reteaching is needed throughout the school experience. (How often have you said to your students, "My goodness, didn't they teach you that in grade X?") So we end up with what we see in so many schools (as verified by NAEP test results): Students in general can do low-level tasks but are universally weak in higher-order work that requires transfer.

The research on learning (considered in greater detail in Chapter 13) merely supports the sobering truth of common sense: If learning is to endure in a flexible, adaptable way for future use, coverage cannot work. It leaves us with only easily confused or easily forgotten facts, definitions, and formulas to plug into rigid questions that look just like the ones covered. Furthermore, we have thereby made it far more difficult for students to learn the "same" things in more sophisticated and fluent ways later. They will be completely puzzled by and often resistant to the need to rethink earlier knowledge. In short, as Lee

Shulman, president of the Carnegie Center for the Advancement of Teaching, put it so well, conventional teaching abets the three "pathologies of mislearning: we forget, we don't understand that we misunderstand, and we are unable to use what we learned. I have dubbed these conditions amnesia, fantasia, and inertia" (Shulman, 1999, p. 12).

Our analysis thus far suggests, then, the need for three types of "uncoverage" in designing and teaching for understanding to avoid forgetfulness, misconception, and lack of transfer:

- Uncovering students' potential misunderstandings (through focused questions, feedback, diagnostic assessment)
- Uncovering the questions, issues, assumptions, and gray areas lurking underneath the black and white of surface accounts
- Uncovering the core ideas at the heart of understanding a subject, ideas that are not obvious—and perhaps are counterintuitive or baffling—to the novice

The evidence of understanding

> What differentiates revolutionary thinkers from non-revolutionary ones is almost never a greater knowledge of the facts. Darwin knew far less about the various species he collected on the *Beagle* voyage than did experts back in England who classified these organisms for him. Yet expert after expert missed the revolutionary significance of what Darwin had collected. Darwin, who knew less, somehow understood more.
>
> —Frank J. Sulloway, *Born to Rebel*, 1996, p. 20

If understanding is about making meaning of facts and transferring knowledge to other problems, tasks, and domains, what does such understanding (or lack of it) look like? What should we be seeing if our students are getting better at understanding what they are learning? To pose this question is to shift from talking about our aims to talking about the evidence of whether our aims have been met.

The Sulloway comment about Darwin suggests one line of inquiry. Consider the words we use in describing understanding at the highest levels of research. We often describe understanding as "deep" or "in depth" as opposed to superficial knowledge. You have to "dig" below the "surface" (i.e., the "cover") to "uncover" unobvious "core" insights. Understanding "takes time and practice." Understandings are "hard won," not immediate—maybe even overlooked or unseen by those with lots of knowledge, as Sulloway suggests. The emphasis in all these connotations is on getting below the surface, to the hidden gems of insight. We cannot *cover* concepts and expect them thereby to be understood; we have to *uncover* their value—the fact that concepts are the results of inquiry and argument.

Notice, then, the difference in the two questions at the heart of grappling with goals related to understanding (and all educational goals more generally) via backward design—the questions for the first two of the three stages:

Stage 1: What should students come away understanding?

Stage 2: What will count as evidence of that understanding?

The second question actually encompasses distinct questions that make up the second stage of backward design:

- Where should we look for evidence? What is the *type* of student work we need to see done well, given the stated standard?
- What should we look for specifically in student performance, regardless of the particular approach, for us to judge the degree to which the student understands?

Loosely speaking, the first question about the evidence involves a design standard for assessment of the work (i.e., what are valid tasks, tests, observations?), and the second question about the evidence concerns the actual evaluation of the work produced, via rubrics or other criteria-related guidelines.

The argument for backward design is predicated on the view that we are not likely to achieve our target of understanding—however we define the term—unless we are clear about what counts as *evidence* of that understanding. And the more we ask that nitty-gritty assessment question, the more many teachers come to understand that they may not have adequately understood understanding.

> ### ■ MISCONCEPTION ALERT!
>
> A *standard* is different from a *performance indicator.* A standard represents a goal and belongs in Stage 1. A performance indicator, such as those found often in bulleted lists under state content standards, represents possible assessment evidence. Making matters more confusing, sometimes the standards also refer to learning activities like those we would put in Stage 3. (See *standard* in the Glossary.)

Why might we be unsure about what constitutes good evidence of understanding? Because the evidence we tend to focus on or that stands out more readily can easily mislead us if we are not careful. When students provide the answer we seek, it is easy to conflate such recall with understanding. Bloom and his colleagues (1956) remind us of the distinction when they recount a famous story about John Dewey:

> *Almost everyone has had the experience of being unable to answer a question involving recall when the question is stated in one form, and then having little difficulty . . . when the question is stated in another form. This is well illustrated by John Dewey's story in which he asked a class, "What would you find if you dug a hole in the earth?" Getting no response, he repeated the question; again he obtained nothing but silence. The teacher chided Dr. Dewey, "You're asking the wrong question." Turning to the class, she asked, "What is the state of the center of the earth?" The class replied in unison, "Igneous fusion." (p. 29)*

The story beautifully illustrates the need to distinguish the content goal from the evidence, as well as the need to stress transferability in the requirements for evidence. Children cannot be said to understand their own answer, even though it is correct, if they can only answer a question phrased just so. Furthermore, they will not be able to use what they "know" on any test or challenge that frames the same question differently, as apparently happened in the state tests mentioned earlier.

Getting evidence of understanding means crafting assessments to evoke transferability: finding out if students can take their learning and use it wisely, flexibly, creatively. The authors of the Taxonomy note, for example, that "real" knowledge involves using learning in new ways. They call this "intellectual ability" and distinguish it from "knowledge" based on recall and scripted use.

Evidence of understanding requires that we test quite differently, then. We need to see evidence of students' ability to "extract" understandings and apply them in situated problems, in performance—something quite different from merely seeing if they can recall and plug in the underlying principles the teacher or textbook gave them.

This requires us to anchor our assessments in prototypical performances in each area, success at which indicates understanding; for example, the ability to design a science experiment, debug it, and revise it in order to determine the chemical content of a substance; the ability to use the facts and skills learned in history to write a credible narrative about a period in local history. We need to see if students with understandably limited ability can nonetheless transfer—that is, recognize what in their repertoire *might* be useful *here,* in this novel situation, and use it effectively. Thus, we would use far fewer narrow prompts that are intended to elicit the "correct" answer to a familiar question.

The "igneous fusion" example is extreme, but the problem strikes home more than most of us may see or care to admit. We are often too ready to attribute understanding when we see correct and intelligent-sounding answers on our own tests. What may trip us up more than we realize is *apparent* understanding, in other words. And that difficulty is likely exacerbated in a world of high-stakes testing and grading. For as long as education promotes a cat-and-mouse game whereby students have incentive to both please us and *appear* to understand what they are supposed to learn (irrespective of whether they do or not), the challenge of assessing for real understanding becomes greater.

In short, we must be careful: It doesn't matter how we term the difference between knowing and understanding as long as we safeguard the real difference. What we call *understanding* is not a matter of mere semantics. It is a matter of conceptual clarity whereby we distinguish between a borrowed expert opinion and an internalized flexible idea. If our assessments are too superficial and fact-centered, we may miss the distinction in the evidence we collect. It does not matter in the end what we call understanding-related targets, but it matters greatly that we safeguard the distinction between "understand" and "know the right answer when prompted." What matters is that we grasp the challenge of assessing for transfer.

Student *mis*understanding and what we can learn from it

Somehow, well-intentioned, able, and attentive students can take away lessons that we never intended. What are we complaining about when we say of students, "They know all the facts, but they put them together all wrong" or, "They just aren't thinking about what they are saying"? *The Catcher in the Rye* is a fixture of high school English courses in the United States, for example, yet many students come away believing the book to be about Holden's "excellent adventure" (to borrow from a recent movie title), the larklike days in the life of a hooky-playing prep school student. Somehow, the fact that Holden is in great emotional pain—and tells the story from his psychiatric hospital bed—is unseen by many students. Similarly, in mathematics, many elementary students struggle mightily with the multiplication of fractions, given the oddity of the answers being smaller than the numbers they started with. Or consider the great challenge of reading: Simple decoding is not so simple. We pronounce "lose" as "loze" and the teacher tells us we are mistaken. But we thought we understood the rule! Why *isn't* the pronunciation of "lose" consistent with the long-vowel rule about words that end in a consonant and *e* (e.g., close, doze, home)?

Misunderstanding is not ignorance, therefore. It is the mapping of a working idea in a plausible but incorrect way in a new situation. Here are some examples:

- An elementary teacher reported the irritation of one of her 4th grade students at not ever seeing lines of longitude and latitude as she flew cross-country with her family.
- A very bright and learned boy, with advanced placement science courses in his background, thought "error" in science was a function of avoidable mistakes, rather than a principle inherent in the enterprise of induction.

Paradoxically, you have to have knowledge and the ability to transfer in order to misunderstand things.

Thus evidence of misunderstanding is incredibly valuable to teachers, not a mere mistake to be corrected. It signifies an attempted and plausible but unsuccessful transfer. The challenge is to reward the try without reinforcing the mistake or dampening future transfer attempts. In fact, many teachers not only fail to see the value in the feedback of student misunderstanding, they are somewhat threatened or irritated by it. A teacher who loses patience with students who don't "get" the lesson is, ironically, failing to understand—the Expert Blind Spot again. For *attentive* students not to "get it" is to show us that what we thought was clear was really not so. For some teachers, perpetual student misunderstanding is therefore threatening, understandably, because it seems to call into question our methods and implied goals. What the naïve teacher may be overlooking, of course, is that the big ideas are rarely obvious. Indeed, they are often counterintuitive, as we noted in Chapter 1. A word to the wise, then: If you hear yourself saying to a class, "But it's so obvious!" you are most likely falling prey to the Expert Blind Spot! Take time to ponder: Hmmm, what is not obvious to the novices here? What am I taking for granted that is easily misunderstood? Why did they draw the conclusion they did?

Making the matter of greater urgency is the fact that research over the past 20 years confirms the surprising depth and breadth of the phenomenon. Many students, even the best and most advanced, can *seem* to understand their work (as revealed by tests and in-class discussion) only to later reveal significant misunderstanding of what they "learned" when follow-up questions to probe understanding are asked or application of learning is required.

As Gardner (1991) explains in summing up the research,

> *[What] an extensive research literature now documents is that an ordinary degree of understanding is routinely missing in many, perhaps most students. It is reasonable to expect a college student to be able to apply in new context a law of physics, or a proof in geometry, or the concept in history of which she has just demonstrated acceptable mastery in her class. If, when the circumstances of testing are slightly altered, the sought-after competence can no longer be documented, then understanding—in any reasonable sense of the term—has simply not been achieved. (p. 6)*

For more than a decade in physics, specific tests have been developed and used as assessments targeting key misconceptions. The most widely used test, the Force Concept Inventory, provides a pre- and post-test instrument for measuring progress in overcoming the most common (and surprisingly persistent) misconceptions.

AAAS, in its *Benchmarks* (1993) and *Atlas of Science Literacy* (2001), has provided a rich account of desired understandings in the sciences, coupled with key misunderstandings connected with them:

> *Comparison of data from two groups should involve comparing both their middles and the spreads around them.*
>
> *The middle of a data distribution may be misleading—when the data are not distributed symmetrically, or when there are extreme high or low values, or when the distribution is not reasonably smooth.*
>
> • The concept of the mean is quite difficult for students of all ages to understand even after years of formal instruction. . . . Research suggests that a good notion of "representativeness" may be a prerequisite to grasping the definitions of mean, median and mode. . . . Premature introduction of the algorithm for computing the mean divorced from a meaningful context may block students from understanding what averages are. (AAAS, 2001, pp. 122–123)

To see how easy it is to misunderstand things we think we all know, consider this more basic science question: Why is it colder in winter and warmer in summer? Just about every student in the United States has been taught basic astronomy. We "know" that the Earth travels around the sun, that the orbit is elliptical, and that the Earth tilts at about 20 degrees off its north-south axis. But when graduating Harvard seniors were asked the question (as documented in a video on the misunderstanding phenomenon produced by the Harvard-Smithsonian Center for Astrophysics), few could correctly explain why (Schneps, 1994).[2] They either had no adequate explanation for what they claimed to know or they provided a plausible but erroneous view (such as, the weather changes are due to the earth being closer or farther from the sun).

The recognition of inevitable learner misunderstanding in even the best minds, in disciplines as seemingly straightforward and logical as science and mathematics, is actually quite old. Plato's dialogues vividly portray the interplay between the quest for understanding and the habits of mind and misconceptions that may be subconsciously shaping or inhibiting our thinking. Francis Bacon (1620/1960) provided a sobering account of the misunderstandings unwittingly introduced by our own intellectual tendencies operating unawares in the *Organon* 400 years ago. He noted that we project categories, assumptions, rules, priorities, attitudes, and matters of style onto our "reality" and then develop countless ways of "proving" our instinctive ideas to be true: "The human understanding . . . when it has once adopted an opinion draws all things else to support and agree with it" (pp. 45–49). Philosophers and psychologists from Kant and Wittgenstein to Piaget and other modern cognitive researchers have attempted to figure out the puzzle of persistent misunderstanding and the naïve conviction that typically accompanies it—and the self-assessment and self-discipline needed to move beyond both.

Practically speaking, we must begin to design assessments in recognition of the need for conceptual benchmarks, not just performance abilities. We need to design assessments mindful of not only the big ideas but also the *likelihood* that those ideas will be misconceived—and will resist being overcome, as in this biology example cited by Shulman (1999):

> *Biology teachers must wrestle with the durability of student misconceptions of evolution and natural selection. Most students in courses that emphasize evolution and natural selection enter these courses as intuitive Lamarckians. They are convinced that any characteristics acquired by one generation are then transmitted to the next generation. The formal instruction emphasizes the Darwinian refutation of that position. These students may earn A's and B's in the course, demonstrating that they now understand the Darwinian perspective, but quiz them three months later and they're once again dedicated intuitive Lamarckians—as indeed are many of the rest of us. I suspect that forms of fantasia are endemic among students and graduates of higher education, many lying in wait for years before manifesting themselves at critical moments. (p. 12)*

Here are some examples of common misunderstandings for some important ideas, and understandings that reflect the overcoming of them:

- *Science is about finding causes.* Scientists find correlations; talk of "causes" is viewed as too philosophical and unscientific. Modern science, economics, and medicine search for statistical patterns. That's why asking "What caused it?" is not necessarily a question doctors can answer, even as they prescribe effective medicines.
- *History is about the facts, what happened.* A historian is a storyteller, not a mere gatherer and purveyor of facts. Why, then, do so few students realize that there can be and are very different stories of the same important history?
- *Light is light and dark is dark.* Not true. Two light beams intersecting at crest and trough can cancel each other out and cause darkness! Noise-canceling headphones use sound to produce silence. Similiarly, mirror-image waves of light or sound cancel each other out.
- *Negative and imaginary numbers are unreal.* Negative and imaginary numbers are no less and no more real than ordinary numbers. They exist to provide the symmetry and continuity needed for essential arithmetic and algebraic laws.
- *Evolution is a controversial idea.* No, the theory of natural selection as the engine of evolution is what is controversial. Theories of evolution predated Darwin by centuries and were not seen as being in conflict with religious doctrine.
- *Irony is coincidence.* Irony is not mere coincidence, though almost every sportscaster misuses the word! Irony is what the wiser person sees that another seemingly wise person does not. The audience sees what Oedipus does not, and the tension between the latter's pride and what we know is the truth is the source of the drama's power.

Given the likelihood of deeply rooted misconceptions and the potential for misunderstanding, a proactive and, for most of us, unfamiliar approach to assessment design is required. To successfully engineer understanding, we have to think backward: What does understanding look like when it is there or not there? We have to be able to describe what it looks like, how it manifests itself, how apparent understanding (or misunderstanding) differs from genuine understanding, which misunderstandings are most likely to arise (thus interfering with our goal), and whether we are making headway in ferreting out and eradicating the key impediments to future understanding. In other words, we have to think through our assessments before we think through our teaching and learning.

Chapter 7

Thinking like an Assessor

> We recognize understanding through a flexible performance. . . . Understanding shows its face when people can think and act flexibly around what they know. In contrast, when a learner cannot go beyond rote and routine thought and action, this signals lack of understanding. . . . To understand means to be able to perform flexibly.
>
> —David Perkins, "What Is Understanding?" in Martha Stone Wiske, Ed., *Teaching for Understanding*, 1998, p. 42

Nowhere does the backward design process depart more from conventional practice than at this stage. Instead of moving from target to teaching, we ask, What would count as evidence of successful learning? Before we plan the activities, our question must first be, What assessment of the desired results logically follows Stage 1? And, specifically, what counts as evidence of the understanding sought?

Three basic questions

Thinking like an assessor boils down to a few basic questions. The first question is *What kinds of evidence do we need* to find hallmarks of our goals, including that of understanding? Before we design a particular test or task, it's important to consider the general types of performances that are implied. For example, regardless of content, understanding is often revealed through the exercises of comparing and contrasting or summarizing key ideas. After mapping a general approach to assessment, we then develop the assessment particulars.

The second question assumes that some particular task has been developed, about which we then ask, *What specific characteristics in student responses, products, or performances should we examine* to determine the extent to which the desired results were achieved? This is where criteria, rubrics, and exemplars come into play.

The third question has to do with a test for validity and reliability of the assessment: *Does the proposed evidence enable us to infer a student's knowledge, skill, or understanding?* In other words, does the evidence (Stage 2) align with our goals (Stage 1), and are the results sufficiently unambiguous? Few teachers are in the habit of testing their designs once the assessments have been fleshed out, but such self-testing is key to better results and to fairness.

In this chapter, we consider the first of the three aspects of thinking like an assessor: considering, in general terms, the kind of evidence needed to assess a variety of learning goals generally and understanding specifically. In the following chapter, we address the other two questions, related to criteria and the issues of validity and reliability.

Figure 7.3
Two Approaches to Thinking About Assessment

When thinking like an assessor, we ask—	When thinking like an activity designer (only), we ask—
• What would be sufficient and revealing evidence of understanding?	• What would be fun and interesting activities on this topic?
• Given the goals, what performance tasks must anchor the unit and focus the instructional work?	• What projects might students wish to do on this topic?
• What are the different types of evidence required by Stage 1 desired results?	• What tests should I give, based on the content I taught?
• Against what criteria will we appropriately consider work and assess levels of quality?	• How will I give students a grade (and justify it to their parents)?
• Did the assessments reveal and distinguish those who really understood from those who only seemed to? Am I clear on the reasons behind learner mistakes?	• How well did the activities work? • How did students do on the test?

General guidelines

We can sum up the concerns in Chapters 7 and 8 by offering the following questions and guidelines to consider when constructing a balanced set of local assessments of understanding:

1. The needed evidence is inherently less direct and more complicated than that obtained from objective tests to assess knowledge and skill. We need to look at more than just the percentage of correct answers. Why? Sometimes getting the right answer occurs as a result of rote recall, good test-taking skills, or lucky guessing. In assessing for understanding, we need to ferret out the reasons behind the answers and what meaning the learner makes of the results.

2. Assessment of understanding requires evidence of "application" in performance or products, but that complicates judging results. What do we do when parts of a complex performance are shaky, but we discern clear insight in the content? Or the result is fine, yet we sense that little insight was required to complete the project? How do we design performances that enable us to make precise judgments about the different parts of performance?

3. Since understanding involves the six facets, do some facets take precedence over others? *Which* performances matter most, in *what* situations? What can we infer, for instance, when the "application" and "explanation" of strategy is strong but the "interpretation" of the situation is weak? Or the particular "application" was ineffective, but verbal analysis and self-assessment makes clear that the learner has a solid understanding of the content and process?

4. Try to have parallel versions of the same content across different assessment formats. In other words, counteract the "messiness" of a complex task with a simple quiz in the same content. Or use constructed response questions on the same content to make sure that correct answers cannot hide lack of understanding. Whenever possible, have parallel assessments in diverse formats improve the quality of the evidence of desired results.

5. Try to anticipate key misunderstandings and develop quick preassessments and postassessments to find out if those misunderstandings were overcome—regardless of what other assessment tasks you are using. For example, the following quick assessment task reveals whether students understand the process of isolating variables as part of a science investigation:

> *Roland wants to decide which of two spot removers is best. First, he tried Spot Remover A on a T-shirt that had fruit stains and chocolate stains. Next, he tried Spot Remover B on jeans that had grass stains and rust stains. Then he compared the results. Is there a problem with Roland's plan that will make it hard for him to know which spot remover is best? Explain.*

6. Given that a single application or product may or may not link to larger goals, regularly ask students to "show their work," give reasons for answers, and show connections to larger principles or ideas in the answers.

7. Given that an articulate explanation may be more a function of verbal ability and verbal knowledge with no real understanding, ask the student to "transfer" that explanation to a new or different problem, situation, or issue.

8. Tap into various facets to broaden the evidence: When demanding a hands-on application (Facet 3), also require interpretation (Facet 2), and self-assessment (Facet 6) to make sure that the final product is not overvalued. Require a blend of perspective and empathy whenever possible.

Figure 11.1
Entry Points for the Design Process