Ninth Edition

Classroom Assessment

What Teachers Need to Know

W. James Popham
Professor Emeritus
University of California, Los Angeles

Director and Publisher: Kevin M. Davis
Content Producer: Janelle Rogers
Development Editor: Curtis B. Vickers
Media Project Manager: Lauren Carlson
Portfolio Management Assistant: Maria Feliberty
Executive Field Marketing Manager: Krista Clark
Executive Product Marketing Manager: Christopher Barry
Operations Specialist: Carol Melville
Cover Designer: Carie Keller
Cover Image: Maskot/Shutterstock/Offset
Full Service Project Management: Thistle Hill Publishing Services, LLC
Composition: SPi Global
Text Font: 9.5/13 Palatino LT Pro

Library of Congress Cataloging-in-Publication Data
Names: Popham, W. James, author.
Title: Classroom assessment : what teachers need to know / W. James Popham, Professor Emeritus, University of California, Los Angeles.
Description: Ninth edition. | New York, NY : Pearson, [2020] | Includes bibliographical references and index.
Identifiers: LCCN 2019013736| ISBN 9780135569108 (pbk.) | ISBN 0135569109 (pbk.)
Subjects: LCSH: Educational tests and measurements. | Examinations—Design and construction. | Examinations—Validity.
Classification: LCC LB3051 .P6142 2020 | DDC 371.26—dc23
LC record available at https://lccn.loc.gov/2019013736

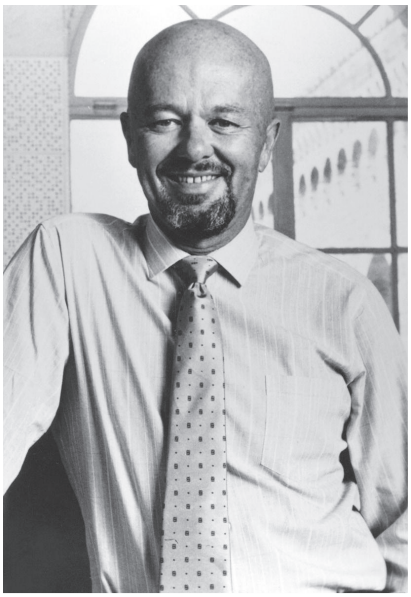
1 19

Pearson

ISBN-13: 978-0-13-556910-8
ISBN-10: 0-13-556910-9

For Sarah:
Former Teacher, Former Principal,
Current Colleague, Current and Future Spouse

About the Author



W. JAMES POPHAM has spent the bulk of his educational career as a teacher. His first teaching assignment, for example, was in a small eastern Oregon high school where he taught English and social studies while serving as yearbook advisor, class sponsor, and unpaid tennis coach. The recompense meshed well with the quality of his coaching.

Most of Dr. Popham's teaching career took place at UCLA where, for nearly 30 years, he taught courses in instructional methods for prospective teachers as well as courses in evaluation and measurement for graduate students. At UCLA he won several distinguished teaching awards. In January 2000, he was recognized by *UCLA Today* as one of the university's top 20 professors of the twentieth century. (He notes that the twentieth century was a full-length century, unlike the current abbreviated one.) In 1992, he took early retirement from UCLA upon learning that emeritus professors received free parking.

Because at UCLA he was acutely aware of the perishability of professors who failed to publish, he spent his nonteaching hours affixing words to paper. The result: over 30 books, 200 journal articles, 50 research reports, and 175 papers presented before research societies. Although not noted in his official vita, while at UCLA he also authored 1,426 grocery lists.

His most recent books are *Transformative Assessment* (2008); *Instruction That Measures Up* (2009); *Transformative Assessment in Action* (2011); *Mastering Assessment* (2011, Pearson); *Unlearned Lessons* (2009, Harvard Education Press); *Everything School Leaders Need to Know About Assessment* (2010); and *Evaluating America's Teachers: Mission Possible?* (2013, Corwin). He encourages purchase of these books because he regards their semi-annual royalties as psychologically reassuring.

In 1968, Dr. Popham established IOX Assessment Associates, an R&D group that formerly created statewide student achievement tests for a dozen states. He has personally passed all of those tests, largely because of his unlimited access to the tests' answer keys.

Preface

Of Ancient Chinese Curses

Perhaps you have heard the ancient Chinese curse, *May you live in an interesting time!* Perhaps you haven't.

Well, I can definitely tell you where and when I first heard this curse—and how puzzled I was by its meaning. The year was 1961, and I was a rookie assistant professor at San Francisco State College. A campuswide speech was to be delivered by Robert Maynard Hutchins, an educational celebrity of that era. Hutchins was the founder of the Great Books Movement and had been the youngest-ever chancellor of the University of Chicago.

It was a simply marvelous speech—so fine, in fact, that I subsequently obtained an audiotape of the address and played it often in my classes. Hutchins opened his address with the following sentence: "Perhaps you have heard the ancient Chinese curse, '*May you live in an interesting time!*'"

As I indicated, upon hearing Hutchins's first sentence, I was immediately perplexed by the meaning of this "curse" that I'd never heard before. After all, if the time in which one lives is "interesting," this would seem to be a positive—not a negative. What's interesting is typically better than what's dull. But then, as Hutchins continued, he pointed out that an "interesting time" invariably involves *changes.* Indeed, the more profound the changes, the more "interesting" the time. And changes, at least for most of us, cause *discomfort.* We must accommodate to what's new. Routine, experience-honed approaches no longer work. New, "interesting" times simply bristle with uncertainty. Hutchins was warning his audience that education in the United States was entering an era of unprecedented change and, as a consequence, U.S. educators should clearly regard themselves as consummately cursed.

Well, if you look at what's taking place these days in this nation's *educational assessment,* you'll quickly conclude that we are smack in the middle of what is, most certainly, another especially "interesting time." To illustrate, during the several years separating the publication of this ninth edition of *Classroom Assessment* from its immediately previous edition, we have witnessed several significant national changes that, almost certainly, will result in "interesting" alterations in the way much of the nation's classroom assessment takes place. On December 10, 2015, for instance, President Barack Obama signed into law the Every Student Succeeds Act (ESSA) as another federal reauthorization of the Elementary and Secondary Education Act of 1965. ESSA replaced the No Child Left Behind Act (NCLB) that had previously served as a reauthorization of 1965's ESEA. Both NCLB and ESSA contain provisions influencing how today's teachers are likely to use classroom assessments with their students. Granted, those federal laws only *influenced* how classroom testing is likely to take place. Federal education statutes such as ESSA typically impose regulations obliging *state* educational authorities to carry out certain sorts of requirements. Usually, these federally imposed constraints—or, sometimes, incentives—focus only on state-level actions, not on what happens in teachers' classrooms. And yet, more often than not, federal regulations for implementing an education statute lead to a state's adoption of programs or policies that soon arrive in the classroom.

As you will see, this is where the "interestingness" of today's classroom assessment situation becomes all too apparent. For starters, according

to U.S. Department of Education press releases about ESSA, our current reauthorization of ESEA is seen as a far less "prescriptive" law than NCLB, the reauthorization replaced by ESSA. In other words, state authorities will have many more choices for carrying out the less restrictive requirements of ESSA than they had with NCLB. More choices, of course, mean more differences in how states attempt to comply with altered federal requirements. Ever since 1965, when ESEA was first enacted as what was essentially a civil rights law, some heavy-duty federal funds have been distributed annually to those states that carry out ESEA's stipulations. Not surprisingly, therefore, authorities in most states try to play the education game in accordance with federal rules so that their state can rake in its share of federal largesse.

And now, finally, we see that although NCLB laid out some pretty constraining obligations for states with respect to their state assessment programs, and monitored those state programs through state-specific review teams who appraised each state's assessment programs, the enactment of a less prescriptive ESSA will often lead to more changes, and more differences, in how a state's teachers are supposed to carry out their classroom assessment endeavors.

If Robert Maynard Hutchins (1899–1977) were still with us, his curse for today's teachers might well be *"May your classroom assessments be interesting!"* Whereas the federal testing regulations associated with NCLB were quite constraining—and offered few opportunities for state-sired innovations, our current incarnation of ESEA, namely, ESSA, encourages more diverse ways to measure students' status.

Teachers' bread-and-butter testing activities associated with classroom assessment—the activities not influenced by federal statutory shifts—remain relatively constant. And those main-line assessment decisions are the ones addressed in much of this oft-revised classroom assessment textbook—the one you'll soon read. A thorough review of the previous edition of the book has taken place for purposes of updating. For example, since the publication of the eighth edition of *Classroom Assessment*, we have seen a substantial increase in the use of computer-*adaptive* testing. This topic, its advantages, and its drawbacks are considered in the new edition. Similar updating has been applied throughout.

Thus, if you are currently a teacher, or are preparing to become one, you will find in the ninth edition of *Classroom Assessment* a complete set of concepts and procedures that will enable you to do a solid job of testing your own students. But, during the last decade or two, we have seen a growing need for educators *and others* to become more knowledgeable about educational assessment itself. It is this advocacy that has led to the inclusion of exciting new features in this new edition of *Classroom Assessment*.

New to This Edition

- **Extensive discussion of the *Standards for Educational and Psychological Testing* and their impact on classroom assessment:** In mid-2014, three of our nation's organizations most concerned with educational measurement released a genuinely significant document—namely, the *Standards for Educational and Psychological Testing*. The standards (guidelines) included therein describe how education tests should be constructed, evaluated, and used. Because the standards in that publication are often relied on to resolve assessment-related courtroom litigation, the tenets set forth will have a profound impact on the way U.S. educational testing takes place in the coming years. The most important of these new standards have been incorporated into all relevant sections of this ninth edition of *Classroom Assessment.*
- **Keys for evaluating instructionally diagnostic testing:** Given today's pervasive pressure to boost students' test scores, a number of test publishers are hawking what they characterize

as "diagnostic tests." To help teachers understand what's necessary for a diagnostic test to make a meaningful contribution to instruction, an important section on *instructionally diagnostic testing* will be found in Chapter 13. Readers who understand what's needed in a test that is truly diagnostic will be better able to choose among an ever-expanding array of supposedly diagnostic assessments now available commercially. Many of today's allegedly useful diagnostic tests should, regrettably, be sent to a shredder. The section on instructionally diagnostic tests will help readers evaluate the merits of commercial or locally developed "diagnostic" assessments.

- **Updated coverage of relevant legislation and policy changes, including the implications of the Every Student Succeeds Act (ESSA) on classroom teachers:** The ninth edition includes extensive coverage of the legislation and policy changes that affect the classroom teacher. Chief among these is the Every Student Succeeds Act (ESSA) passed in 2015.
- **Expanded treatment of technology and its implications for classroom assessment:** The ninth edition has responded to the rising popularity of computer-based tests, particularly computer-adaptive tests, by dedicating considerable coverage to their methods and their implications for classroom assessment.
- **Reliability, validity, and fairness in action:** Throughout the ninth edition, care has been taken to provide concrete examples of how assessment's so-called Big Three—reliability, validity, and fairness—come into play for classroom teachers in a variety of instructional settings.
- **Parent Talk:** Considerable attention is paid in this ninth edition to an important audience for teachers—namely, *parents*. Teachers often need to explain to parents such assessment-related topics as why a child scored a certain way on a standardized test or how a teacher's classroom exams are related to the grades a child receives. So, in every chapter, you'll find a feature titled *Parent Talk*. In these Parent Talk segments, a situation is described in which a teacher needs to explain something about assessment to parents. I next indicate what I would have said to the parents if I had been the teacher in that situation. What I hope readers will do, then, is decide how *they* would have responded had they been placed in this same situation. In fact, readers might even say *aloud* (in private, if they have any sense) what they'd tell parents. Teachers who can talk sensibly to parents about assessment-related concerns will find they're able to establish more effective rapport with parents. Such teachers will get along far better with parents than will teachers who convey to parents the idea that assessment is an exotic measurement mystery, well beyond the perceptive powers of mere parents.
- **MyLab Education:** One of the most visible changes in the ninth edition, and also one of the most significant, is the expansion of the digital learning and assessment resources embedded in the eText and the inclusion of MyLab Education in the text. MyLab Education is an online homework, tutorial, and assessment program designed to work with the text to engage learners and improve learning. Within its structured environment, learners see key concepts demonstrated through real classroom video footage, practice what they learn, test their understanding, and receive feedback to guide their learning and ensure their mastery of key learning outcomes. Designed to bring learners more directly into the world of K–12 classrooms and to help them see the real and powerful impact of classroom assessment concepts covered in this book, the online resources in MyLab Education include:

  Self-Checks. In each chapter, self-check quizzes help assess how well learners

have mastered the content. The self-checks are made up of self-grading multiple-choice items that not only provide feedback on whether questions are answered correctly or incorrectly, but also provide rationales for both correct and incorrect answers.

Application Exercises. These exercises give learners opportunities to practice applying the content and strategies from the chapters. The questions in these exercises are usually constructed-response questions. Once learners provide their own answers to the questions, they receive feedback in the form of model answers written by experts.

Podcasts. In the podcasts, readers will learn about real events related to the key concepts in each of the 16 chapters in the book. These stories help awaken readers to the real-life consequences and power of assessment done well and done poorly.

- **Testing Takeaways:** To help promote assessment literacy, the ninth edition of *Classroom Assessment* has introduced a new feature called Testing Takeaways. Each Testing Takeaway is an easily digestible one-page explanation of a key assessment concept. These features are designed to be shared with parents, legislators, and any other person for whom assessment and assessment policy have bearing. An expanded discussion of the importance of Testing Takeaways and their application follows.

Promote Assessment Literacy: Share a Testing Takeaway

Because test results influence what happens not only in classrooms but also in state legislatures, it's increasingly important for educators, parents, students, and lawmakers to become more assessment literate. But what is this thing called "assessment literacy" that you're being urged to promote?

Assessment literacy consists of an individual's understanding of the fundamental assessment concepts and procedures deemed likely to influence educational decisions. Becoming assessment literate does not require someone to magically morph into a sophisticated testing wizard. Rather, an assessment-literate educator needs only to grasp a modest number of essential understandings—those that are apt to have a real-world impact on the education of real-world children.

To help promote assessment literacy, the ninth edition of *Classroom Assessment* has introduced a new feature, Testing Takeaways, which are a set of easily digestible explanations. Candidly, these new Testing Takeaways were not written for you, the reader, but instead are intended to be read by persons or groups with whom you decide such takeaways should be shared. By doing so, you are helping *others* acquire a few useful assessment insights that may, in time and with greater study, blossom into full-on assessment literacy.

There is a definite rationale for including this new feature in the ninth edition of *Classroom Assessment*. All of the Testing Takeaways are focused on an aspect of assessment literacy. The rationale supporting the inclusion of these new segments stems from a belief that increased assessment literacy—on the part of educators and other educational stakeholders—constitutes *the most cost-effective way* to improve our schools. Put simply, those who understand the basics of educational testing will make fewer mistakes than those who possess scant knowledge about educational assessment. Moreover, in contrast to other, patently effective but costly, school-improvement strategies (such as substantially raising teachers' salaries or reducing class sizes to permit more tailored per-student instruction), the *costs* necessary for promoting assessment literacy are truly trifling.

By the time you have finished reading this book, there'll be no doubt that you will, yourself, have become assessment literate. Unfortunately, not every educator will read this book. As the book's author, I am personally dismayed by this

likelihood. In truth, my dismay is only *partially* due to the contemplation of the non-existent book royalties.

Yet, if we make the reasonable assumption that assessment-literate educators will make fewer test-related mistakes than educators who are not assessment literate, then we surely need to promote assessment literacy among the nation's educators. And not only among educators! Other groups who could benefit from a better understanding of the basics of educational testing include educational policymakers such as members of local or state school boards, parents of school-age children, citizens in general, and even students themselves. Obviously, it might be necessary to adjust certain assessment fundamentals to the concerns of different groups. Yet, most of the basic assessment understandings needed by teachers and school administrators will be useful to other stakeholders as well.

Thanks to Pearson, the publisher of the book you're about to read, in the following pages you will find a set of 16 *Testing Takeaways,* one per chapter, for which Pearson is relaxing copyright protection and authorizing their use *for increasing assessment literacy.* You have permission to make direct copies of each of these one-page Testing Takeaways (hereinafter abbreviated as TTs). A digitally shareable version is available from www.pearsonmylabandmastering.com.

Each TT has been written so that it represents a meaningful, self-contained chunk of assessment-related information pertinent to one of the ninth edition's 16 chapters. Sometimes a TT addresses the overall thrust of a chapter. In other cases, a TT describes an issue related to, but not identical to, a chapter's main message. Certain of the TTs will surely be of more interest to you than others. All TTs attempt to deal with a dollop of assessment-related content that's worth knowing—sometimes only by educators, sometimes by everyone.

The TTs are located at the very end of each chapter. Ideally, you will look them over so you can decide which of them you regard as worth sharing with others. They are not intended to summarize what's in a chapter, but rather to provide a roughly 400-word mini-essay related to the chapter. The TTs are supplied in the hope that you will use them to kindle an interest in assessment literacy on the part of others. This is why I say that the TTs were not really written for you.

If you are using this book in connection with a course focused on educational assessment, then your instructor may have some specific suggestions regarding the use of the TTs. (Because I have spent roughly half my life as a university professor, I urge you to be particularly deferential to your professor's preferences.) For example, instructors who support the promotion of assessment literacy may find it useful, collaboratively with their students, to discuss which TTs seem to be appropriate (or inappropriate) for different audiences.

However, if you happen to be reading the book on your own (yes, such an occurrence might actually take place!) or are using the book in connection with some sort of professional development activity, here are a few suggestions regarding how *you* might personally promote greater assessment literacy on the part of others by using the 16 TTs written for that specific purpose:

- *Send to a friend or colleague.* If you think coworkers or friends would be interested in (or benefit from) the focus of a given TT, you might fire it off as a digital or hard copy to those individuals. You'll probably need to supply a brief introduction so that recipients don't think they've been surreptitiously singled out as gravely needing remedial assessment information.
- *Relay to local news outlets.* Many community newspapers still deal with education-related issues of potential interest to their readers. Were the editors of such newspapers (or the program directors of local television stations) to receive a digital copy of a TT, along with an explanatory note from you, there might be sufficient interest to warrant occasional inclusion of certain TTs in their offerings.

- *Include in a school's parent newsletters.* Because many parents are understandably concerned about the ways educational assessments affect their children, if you find one or more of the TTs that you believe might be of interest to parents, you could send digital or hard copies home in parental newsletters. If, in a cover letter to parents, you could indicate a willingness to provide a list of parent-friendly introductions to educational testing, this would surely contribute to an expansion of assessment literacy among this significant set of educational stakeholders.
- *Initiate a professional development group dealing with assessment.* All teachers within a school could be provided with one or more of the TTs, along with an invitation to collaborate in some ongoing variant of professional development regarding educational assessment. If some participants in such a group had access to the book you are about to read, those educators could periodically supply assessment-relevant content to the group beyond what is found in the TTs themselves.
- *Feature a TT as a segment of regular faculty meetings.* Perhaps you are a teacher in a school where professional topics are treated during routine faculty meetings. If so, then by soliciting reactions to a list and brief descriptions of the available TTs, you might isolate sufficient interest in certain TTs to warrant future discussions of those topics during upcoming faculty meetings.

Other uses of the TTs, of course, are possible—and may be more effective than the suggestions proffered here. Nonetheless, because increased assessment literacy on the part of educators and other clienteles will surely provide more successful educations for our students, give the 16 TTs a serious look—thinking all the while about which groups might most benefit from reading particular TTs. If you can become a personal emissary promoting increased assessment literacy for one or more groups of potential interest, you'll be doing something commendable for kids.

Will This *Testing* Book Help You *Teach* Better?

Teachers these days who don't recognize that educational assessment impinges on their work are teachers in serious need of impingement therapy. Rarely, indeed, does a day go by in the Monday-through-Friday life of today's teachers when testing does not have an impact on one or more of their classroom decisions. It was not always this way.

Eons ago, I was a high school teacher in eastern Oregon. (It was so long ago that my friends contend Oregon must have been a territory rather than a state.) Way back then, we administered standardized achievement tests in our classes. However, students' scores on those tests made no difference in how we taught. Pressures to raise our students' scores on those achievement exams were nonexistent. We taught pretty much as we saw fit. But, of course, the world of education is different today—much different.

And even before those teaching days, when I was preparing to be a teacher, little attention was given to testing. In truth, the only time my professors actually dealt with educational tests was when, during an educational psychology class, we spent an entire week on the making and massaging of multiple-choice items. My fellow prospective teachers and I were not being prepared for educational assessment because, back then, educational assessment truly did not have an impact on teachers' decision making.

But today, educational tests certainly make a difference in what currently takes place in our classrooms. For openers, today's teachers find themselves directly in the crosshairs of some heavy-duty accountability artillery aimed at evaluating schools *and sometimes teachers* at least in part according to students' scores on accountability tests. A school's staff can be "restructured," or

a school can be completely shut down, if its students don't perform well enough on externally administered accountability exams. Yes, teachers can be tossed. It is a scary time.

Fortunately, during the last two decades, growing numbers of educators have learned that the skillful use of classroom testing can make huge differences in how well students learn. Classroom assessment, if employed *formatively*, can dramatically increase the effectiveness of a teacher's *teaching*. And yet, sadly, we rarely see more than a token use of classroom assessments in the way that research clearly tells us will benefit students.

For both of those reasons, then, every experienced teacher and every teacher-in-training need to master the essentials of educational assessment. And that's why this book was first written and then revised so frequently. Its title, *Classroom Assessment: What Teachers Need to Know*, captures the book's intent. Readers won't be asked to learn any nice-to-know exotics about educational measurement. No, what's contained in this book is the stuff today's teachers *need to know* if they are truly going to be first-rate professionals.

Ancillary Materials

The following resources are available for instructors to download on www.pearsonhighered.com/educators. Instructors can enter the author or title of this book, select this particular edition of the book, and then click on the "Resources" tab to log in and download textbook supplements.

Instructor's Resource Manual with Test Bank (ISBN 0-13-556913-3)

The Instructor's Resource Manual includes suggestions for teaching a classroom assessment course using this text, additional learning activities, guidance on using MyLab Education, and a Test Bank.

PowerPoint Slides (ISBN 0-13-556893-5)

The PowerPoint slides include key concept summaries and other aids to enhance learning. They are designed to help students understand, organize, and remember core concepts and applications.

TestGen (ISBN 0-13-556919-2)

TestGen is a powerful test generator that you install on your computer and use in conjunction with the TestGen test bank file for your text. Assessments, including equations, graphs, and scientific notation, may be created for both print and online testing.

The tests can be downloaded in the following formats:

TestGen test bank file—MAC

TestGen test bank file—PC

Angel TestGen Conversion

Test Bank for Blackboard Learning System

Desire to Learn TestGen Conversion

Moodle TestGen Conversion

Sakai TestGen Conversion

Test Bank for Blackboard CE/Vista

Acknowledgments

To my colleagues who offered recommendations about how to improve this latest edition, I offer my sincere appreciation: Heidi Legg Burross, University of Arizona; Melanie DiLoreto, University of West Florida; Thomas J. Sheeran, Niagara University; and Maria D. Vasquez, Florida Atlantic University.

I am appreciative of the overall guidance in this project of Director and Publisher Kevin Davis, my very own Pearson Poohbah, whose executive preferences I obsequiously and instantly obeyed. He has been overseeing the periodic rebirth of this classroom assessment book for a good many years now, and doing a high-class job of it. I was simply delighted when he accepted my proposal to include this edition's new Testing Takeaways in an explicit attempt to promote increased assessment

literacy among others. If we learn that a modest number of such limited waivers of copyright protection can broaden the reach of a book such as this, there is no reason why such a strategy might not be employed as well for content other than educational assessment. Kevin displayed genuine foresight in agreeing to try out this edition's copyright-relaxation outreach strategy.

When a textbook has been through the editorial re-birthing process as often as this one has, an author becomes particularly attentive to the caliber of the production team that transforms words on paper into a *bona fide* book. The following folks were atypically wonderful in creating this ninth edition, and they have my lasting gratitude: Janelle Rogers, Pearson Content Producer; Lauren Carlson, Pearson Media Producer; Krista Clark, Pearson Marketing Manager; and Andrea Archer and Angela Urquhart, Thistle Hill Publishing Services.

Finally, any writer will tell you that an editor can determine a book's quality and, depending on the editor, determine a writer's sanity. This ninth edition was my first opportunity to work with Curtis Vickers, who, as the designated editor for this edition, went through every word—ladling out his criticisms softly and occasionally granting praise when I chose the appropriate forms of pronouns to properly modify gerunds. He was always prompt and, most of all, immersed himself in my blather so he could suggest more beneficial blathering techniques. This new edition of *Classroom Assessment,* and I myself, are in his debt.

W. J. P.

Brief Contents

Contents

Chapter 1
Why Do Teachers Need to Know About Assessment?

Chief Chapter Outcome

An understanding of why it is that four traditional and three recent reasons for educators to assess students should dispose teachers to learn more about the fundamentals of educational assessment

Learning Outcomes

1.1 Identify key legislation, discuss the history of assessment in the United States, and differentiate between testing and assessment.

1.2 Using the four traditional and three modern reasons for why teachers assess found in Chapter 1, identify and explain the rationale for the benefits of teachers as functional assessors.

Teachers teach students. That hardly constitutes a breakthrough insight. Just as welders weld and plumbers plumb—teachers teach. That's why they're called teachers.

But what is a bit less obvious is that most teachers teach because they *like* to teach. Primary teachers like to teach younger children. High school teachers like to teach older children. Most high school teachers also like to teach about a particular subject. (Have you ever noticed how mathematics teachers' eyes get misty when they introduce their students to the raptures of the Pythagorean theorem?) Yes, most teachers love to teach. It is because they enjoy what they do that they waded through a medley of preservice teacher education courses, conquered the challenges of student teaching, and hopped the myriad hurdles of the certification process. Teachers overcame these obstacles in order to earn annual salaries that, particularly during the first few years, are laughably low. Yes, there's little doubt that teachers groove on teaching.

Although teachers like to *teach,* they rarely like to *test.* Yet here you are—beginning a book about testing. How can I, the author, ever entice you, the reader, to become interested in testing when your heart has already been given to teaching? The answer is really quite straightforward. Teachers who can test well will be better teachers. Effective testing will enhance a teacher's instructional effectiveness. Really!

If you're willing to suspend any preconceptions about testing while you're reading this book, particularly any negative ones, I'll make a pledge to you.

If you tackle this text with even half the enthusiasm you might bring to a teaching assignment, I promise you'll discover how *testing will make you a much better teacher*. And, because I've been a teacher for more than 50 years, it's a promise I'll keep. Teachers should definitely not break promises to teachers. (Teachers' promises to administrators, on the other hand, should be regarded as eminently renegotiable.)

But before the book attempts to convince you, ever so subtly, that testing can be a boon to teaching, you need to get a fix on your own *current* views about educational testing. And, because this is a book about testing, what better way to help you explore those attitudes than to have you take a self-test that was devised just for readers of this book?

So, on the adjacent page (it's on the *right* from where you're currently reading!) you'll find a brief self-test similar to the ones you've surely encountered in many widely read magazines. I saw one such self-test in a health magazine recently. It was entitled "How Long Will You Live? A Self-Test." Frankly, I was afraid to try it. As one grows older, one gets more cautious.

But you have nothing to fear by taking the self-test that's been whipped up for you. To emphasize its brevity, it is entitled "A Terse Self-Test About Testing." It is an example of commonly used attitudinal inventories. Later, in Chapter 10, you'll learn more about attitudinal inventories. But for now, please take a crack at page 6's teensy self-test. The way to interpret your responses is given as a footnote at the bottom of the page.

Echoes of ESEA

Anyone who has completed even an introductory course in U.S. government knows that just as state laws can overturn the regulations enacted by local communities, federal laws can overturn state laws. When it comes to the art of overturning, federal folks clearly have the heftiest muscles. Interestingly, we often see those federal muscles flexed with respect to classroom assessment. The origins of the tests that we see used in today's classrooms can sometimes be traced directly to a state or a federal statute dealing with educational assessment. In the instance of classroom assessment, by far the most potent piece of test-influencing legislation bearing on such testing was the federal Elementary and Secondary Education Act (ESEA) of 1965.

Enacted more than a half-century ago, ESEA was a key component of President Lyndon B. Johnson's "Great Society." This precedent-setting law constituted what—at that time—was the federal government's most serious commitment to providing all the nation's K–12 youth with equal access to high-quality education. In contrast to previous federal spending on public education, ESEA's financial support to the nation's schools was downright gargantuan.

As set forth in the new law, unprecedented funds were doled out to states in support of instructional materials, professional development, resources to support educational programs, and the promotion of parental-involvement initiatives.

But because of the unprecedented magnitude of ESEA's federal funding for schools, key congressional leaders (including Robert F. Kennedy, who was then the junior U.S. senator from New York) insisted that the new legislation contain provisions intended to systematically evaluate whether these new federal dollars were being well spent. Formal educational evaluation was, in truth, birthed almost overnight by the passage of 1965's ESEA.

This potent federal statute has been reauthorized every 5–10 years since its enactment, and those reauthorizations have usually contained alterations, sometimes significant, to the law's key provisions. The various subsections of ESEA are designated as "titles." Title I, for example, established a program to distribute federal funding to districts and schools serving a high proportion of students from low-income families. Title I typically gets the most attention from policymakers because it accounts for roughly 80 percent of the total funding that's authorized under ESEA. Programs supported under Title I, at least at the outset, were intended to reduce the skill gaps in reading, writing, and mathematics between children from low-income households and children from the middle class. In general, this is still the case.

From a *classroom-assessment* perspective (and you will recall that classroom assessment is the focus of this book) what's significant about ESEA is that its many reauthorizations have sometimes led to changes in federal evaluation-related requirements that have a meaningful influence on the assessment practices recommended to states, districts, and schools. You will find several such changes embodied in the most recent incarnation of ESEA, and those modifications will be addressed as you saunter joyfully through this book's fun-filled pages.

The Every Student Succeeds Act (ESSA) was signed into law by President Barack Obama on December 10, 2015, and it succeeded the No Child Left Behind Act (NCLB) that was signed by President George W. Bush on January 8, 2002. If you have even modest calendrical and mathematical skills, you can see that, unlike the typical ESEA-reauthorization period of about 5 years, the reauthorization gap between NCLB and ESSA was more than 13 years.

Some observers suggest that the lengthy delay between the two authorizations was chiefly attributable to a markedly increasing polarization of the two major congressional parties. As was the case with the original ESEA, however, both NCLB and ESSA were passed by Congress with substantial bipartisan support. Whatever the cause for the uncommon delay in ESEA's most recent reauthorization, some of the alterations in the law are apt to have a direct impact on the kinds of directives about classroom testing issued to teachers by state and district education authorities.

ESSA's predecessor—NCLB—had attempted to reduce the growing achievement gap that left poor and minority students in failing schools, but it had also introduced a deliberately stringent accountability structure. As President Bush said during the law's signing ceremony (televised by C-Span), the law's fundamental premise is that "every child can learn, we expect every child to learn, and you must show us whether or not every child is learning." As you can infer, NCLB was laced with a strong *"show us"* orientation.

ESSA attempted to preserve the spirit of NCLB, yet it set out to remedy drawbacks that many critics believed stemmed from its "one size fits all" shortcomings. As he signed the act into law, President Obama asserted that NCLB was intended to promote high standards, close the achievement gap, and increase accountability. He regarded these as appropriate goals but he also believed that, in practice, the law failed to consider the distinctive needs of each community. NCLB, he observed in official press releases, had "led to too much testing during classroom time. It often forced schools and school districts into cookie-cutter reforms that didn't always produce the kinds of results that we wanted to see."

In looking back at the two most recent renditions of 1965's ESEA, then, we see that whereas NCLB and ESSA have fundamentally similar goals, ESSA attempts to provide greater flexibility, so that states and districts can particularize their programs for implementing chief provisions of the current successor to ESEA. The success of a given state in carrying out the provisions of ESSA's various titles, however, is supposed to be more closely controlled at the state level than at the federal level, as was the case in NCLB.

Before we depart from federal influence over the sorts of assessment provisions that have a tangible effect on a teacher's classroom testing procedures, it is important for you to recognize a few fundamentals about the ways in which legislatively enacted statutes can determine how teachers test their students. You see, once a federal bill (or, for that matter, a state bill) has been signed into law, a set of regulations is typically readied to guide the law's implementation. Typically, those regulations are drafted by the appropriate governmental agency (in the instance of a federal education-related statute, this would be the U.S. Department of Education) and then are made available in draft form, for a review period, so that those who wish to comment on the emerging regulations can do so. After the suggestions of commentators are taken into consideration, a final set of regulations is issued. These final regulations function as a way of operationalizing the law itself, which often contains segments that are somewhat ambiguous.

Because the nature of teachers' classroom assessments is often influenced by the way a state implements a law such as ESSA, the U.S. Department of Education calls on a special collection of *peer reviewers* who scrutinize a given state's assessment and accountability intentions regarding the implementation of ESSA. State officials, therefore, submit their state's plans for federal peer review and must often make identified changes in those plans before receiving peer-review approval. A federal peer-review panel of, say, a half-dozen assessment specialists, determines whether a state's accountability approach, including its state and local assessments, are sufficiently in accord with ESSA's stipulations. Failure to secure peer-review approval can result in withholding of federal funds, so peer-review approbation is typically sought with substantial zeal.

Classroom teachers who are uncertain about some of the ESSA-influenced issues to be addressed in the upcoming pages, therefore, will often find it helpful to seek the advice, first of district assessment specialists, and then of state-level authorities (such as the assessment personnel in a state department of education),

about how best to carry out classroom assessment so that it meshes with state or federal regulations. Although the membership of a given state's peer-review panel is typically quite stable, changes in panelists can sometimes lead to subtle reinterpretations of ESSA. It is useful, therefore, for teachers to check periodically to see whether any significant alterations have been made in a given state's ESSA expectations.

COMMON STANDARDS = UNCOMMON RUCKUS

As a consequence of a collaborative effort by the National Governors Association (NGA) and the Council of Chief State School Officers (CCSSO), in 2009 we saw 48 states, 2 territories, and the District of Columbia initiate a project culminating in the creation of the ***Common Core State Standards (CCSS)***. The English language arts (ELA) and mathematics learning outcomes embodied in the *CCSS* were intended, as a group, to represent a more challenging set of curricular expectations than were found at that time in many states' official instructional goals. As set forth in the *CCSS*, the key math and ELA content standards—that is, the sanctioned curricular targets at different grade levels—were identified.

In truth, the actual development of the *CCSS* was carried out in a less than totally transparent manner. This was because information about the nature of the content standards themselves—and the specific developmental procedures being used to identify those content standards—was not readily available to interested onlookers. Indeed, many complaints regarding the "covert" building of the *CCSS* were voiced by frustrated would-be observers of this important curriculum-building initiative.

Transparency notwithstanding, however, the resulting sets of math and ELA curricular targets were generally conceded to constitute high-quality collections of more demanding curricular goals than the official sets of intended learning outcomes then seen in some of the 50 states. Indeed, the high quality of the 2010-released *CCSS* surely contributed to their widespread adoption; by 2011, 45 states and the District of Columbia had signified acceptance of the new standards. In recognition that the ELA and mathematics curricular targets contained in the *CCSS* represented the essential features of what a state's schools should be promoting in its students, the emergence and pervasive adoption of these curricular goals constituted a landmark in the history of American schooling.

Let's look, ever so briefly, at what these curricular aims are—with a definite commitment to return in the next chapter for a closer look at the *CCSS*. In Chapter 2, you will see how the two sets of curricular aims identified in the *CCSS* are organized, as well as learn what some of the developers of those state standards were hoping to accomplish.

Let's be clear about what the two collections of ELA and math content standards are. They represent the curricular outcomes sought for the nation's students—that is, the knowledge and cognitive skills students are supposed to acquire in school. Because the reauthorized version of ESEA that was then in place (NCLB) had allowed each state to select its own curricular aims, its own tests to

A Terse Self-Test About Testing

Directions: For each of the statements below, use the following answer key to indicate how you react to the statement:

SA = *Strongly Agree*
A = *Agree*
U = *Uncertain*
D = *Disagree*
SD = *Strongly Disagree*

There are no right or wrong answers, so please answer frankly by circling the appropriate response for each statement.

1. The chief reason why teachers should give classroom tests is to determine students' grades.	SA	A	U	D	SD
2. Teachers should typically plan instruction that focuses on the skills or knowledge represented by a test.	SA	A	U	D	SD
3. In their classroom tests, teachers should use only items that can be scored objectively.	SA	A	U	D	SD
4. There are other legitimate indicators of a teacher's instructional effectiveness besides students' test scores.	SA	A	U	D	SD
5. A teacher has no business measuring students' confidence in their ability to do schoolwork.	SA	A	U	D	SD
6. Today's nationally standardized achievement tests should never be used to supply evidence about how well teachers are instructing children.	SA	A	U	D	SD
7. Teachers rarely need to determine the reliability of their own classroom tests.	SA	A	U	D	SD
8. It is impossible to judge the quality of students' written compositions with any meaningful accuracy.	SA	A	U	D	SD
9. The enormous pressure to boost students' scores on important tests permits teachers to employ almost any sort of score-improvement preparation activities.	SA	A	U	D	SD
10. Significant classroom tests should typically be built before a teacher plans instruction.	SA	A	U	D	SD

Self-Test Interpretation Guide: For statements 2, 4, 6, 7, and 10, use the following scoring key: SA = 5, A = 4, U = 3, D = 2, and SD = 1. For statements 1, 3, 5, 8, and 9, use the following scoring key: SA = 1, A = 2, U = 3, D = 4, and SD = 5. The highest possible total score is 50; the lowest possible total score is 10. The higher your total score, the more sensible is your view of educational testing. After finishing this book, you might wish to retake this terse self-test (without looking at your earlier answers, of course). If you come up with a postbook score that's substantially *lower* than your prebook score, then we should *both* be worried.

assess students' mastery of those aims, and its own cut-scores (that is, achievement standards) to signify students' mastery of those curricular aims, making sense out of the NCLB-spawned accountability picture in U.S. public schools was almost impossible. In an effort to rectify this chaotic situation, the NGA and CCSSO set out in late 2009 to provide a more suitable set of curricular targets for the nation's schools. The CCSSO is the organization of the state officials, elected or appointed, who head each state's public schools. The NGA performs a comparable function for the nation's governors.

On June 2, 2010, the CCSSO and NGA released the *CCSS* (National Governors Association, 2010). As noted earlier, many states soon accepted these standards—these "expectations for student knowledge and skills that high school graduates need to master to succeed in college and careers" (**www.corestandards.org/assets/ccssi-introduction.pdf**). Given the long-standing reluctance of state education officials to abandon "local control" over important educational decisions such as curricular outcomes for students, the widespread adoption of the *CCSS* was genuinely astonishing. In essentially a single year, the CCSSO and NGA had crafted sets of national mathematics and ELA curricular aims defensible enough that all but a few states soon hopped aboard the *CCSS* Express.

The widespread and remarkably rapid adoption of the *CCSS* by so many states, however, did not take place merely because of the merits of a more carefully crafted set of challenging curricular targets. The role of nongovernmental organizations in nurturing such significant changes in U.S. education is now better understood. In the June 7, 2014, issue of the *Washington Post*, Lyndsey Layton reports that a major player in the adoption of the *CCSS* was the Bill and Melinda Gates Foundation. In an article entitled "How Bill Gates Pulled Off the Swift Common Core Revolution," Layton reveals that the Gates Foundation supplied more than $200 million not only for the actual development of the *CCSS* itself but also for building political support across the nation—often convincing state officials to make systematic and expensive changes in their curricular aspirations. Moreover, the foundation spread funds across the entire political spectrum, distributing dollars galore to the two major U.S. teachers' unions and such business groups as the U.S. Chamber of Commerce—organizations that have historically clashed but that soon became outspoken proponents of the Common Core. As Layton reports, within 2 years of the Gates Foundation's decision to support the Common Core, 45 states and the District of Columbia had fully endorsed the *CCSS*.

But the curricular aims embodied in the *CCSS* were destined to serve as much more than lofty statements of curricular intent that, like so many previously crafted sets of curricular aims, typically languished in rarely read reports. This is because, soon after the release of the *CCSS* in mid-2010, the federal government announced its intention to fund one or more consortia of states whose mission it would be to create assessments suitable for measuring students' mastery of the skills and knowledge embodied in the *CCSS*. Two such assessment consortia were selected by federal authorities (from competing bidders) and were funded with approximately $175 million each to create assessments that, by the 2014–15 school

year, could be used to determine students' mastery of the *CCSS*. The two consortia were the *Partnership for the Assessment of Readiness for College and Careers (PARCC)* and the *Smarter Balanced Assessment Consortium (SBAC)*. Each of these consortia was initially composed of about 20 to 25 states, all of which agreed to promote students' mastery of the curricular goals represented by the *CCSS*.

It should be clear that the nature of the assessments devised by PARCC and SBAC would most likely have a considerable impact on America's public schools. Because the curricular aims being pursued by so many states would be identical, and the assessments used in those states would also be identical, comparisons among states' student performances could now be carried out in ways that heretofore had been impossible. The evaluative impact of such evidence, of course, was apt to be substantial.

As the assessments created by the two consortia became more widely understood, straightforward comparisons among states—the comparisons originally foreseen by proponents of the two assessment consortia—grew less likely to take place. Not only were the reporting categories and the cut-scores used by the two consortia dissimilar, but states were allowed unanticipated degrees of local determination in what was taught and what was tested. The original aspiration of the CCSSO and the NGA to install common curricular targets, mastery of which could be measured by essentially identical or statistically equivalent tests, was not to be.

In 2018, for example, although 35 states still regarded the *CCSS* as their official state curricular goals, many states had made minor changes in those curricular goals and, accordingly, described them by other labels). Eleven states had announced a major Common Core rewrite or replacement, four states never adopted *CCSS*, and one state adopted only ELA Common Core goals—not mathematics goals. As this version of *Classroom Assessment* was shipped off to the publisher in early 2019, the number of states and the District of Columbia using SBAC or PARRC assessments was 15—5 fewer than in 2016 or 2017.

Given that about a dozen states still use consortia-created assessments, what about the many other states? Where do their annual accountability tests come from? In many ways, we see state policymakers returning to the kinds of assessment operations they had in place before the arrival of the *CCSS* and the two sets of consortium assessments that, in league with the *CCSS*, were thought by some observers as likely to alter forever the nature of state-adopted curricula and annual assessments. During those pre-NCLB and pre-ESSA days, most states devised their own curricular targets in mathematics, ELA, and sometimes science. The chief determiners of those curricular aims were state-selected educators, K–12, and sometimes university content specialists. Then, to assess students' mastery of the state-selected curricula, most states publicly issued a request for proposals (RFP) so that independent vendors such as AIR, ETS, and Pearson Assessment could bid on developing and, if selected, administering a state's annual accountability tests. This appears to be the current direction of state-level annual testing. Even though we sometimes see small groups of states collaborating in such assessment ventures as a way of economizing (especially for tests

covering particular content areas such as science, or for special student groups such as limited-English speakers), many states are essentially going it alone once more in the generation of annual accountability assessments.

The reason why teachers need to maintain at least a nodding familiarity with what's happening in their own state with respect to implementation of the ESEA reauthorization currently in place (in this instance, ESSA) is that the nature of those federally sired, often state-massaged decisions regarding the nature of state assessments can have a substantial impact on the way teachers should be building their own classroom assessments. For instance, if a state's annual ESSA accountability tests in grades 3–8 feature a substantial emphasis on students' responding to at least a considerable number of constructed-response tests (such as short-answer or essay items), then most teachers in that state will understandably make sure that their own teacher-made classroom assessments also give students practice in responding to such constructed-response items. Teachers who test their students exclusively with multiple-choice or binary-choice items are setting their students up for failure on the annual state tests. Yes, knowledge of federal or state assessment practices can sometimes make a whopping difference in the way teachers build their own classroom assessments. Moreover, the manner in which a state's officials decide to annually assess students' attainment of state-approved curricular targets often influences the way teachers in that state plan and implement their own instructional activities.

The *Standards for Educational and Psychological Testing*

The *Standards for Educational and Psychological Testing* (2014), first published in 1966, contains a set of professionally approved expectations for the way educational and psychological tests ought to be built and used. The *Standards* not only include a series of comments regarding the way that educational and psychological tests should be evaluated, but they also lay out a specific series of detailed "standards"—that is, mandates regarding what is appropriate in the nature and use of educational and psychological tests. This significant document is published by the American Educational Research Association (AERA) and is approved by that organization, as well as by the American Psychological Association (APA) and the National Council on Measurement in Education (NCME).

Because the *Standards* are often invoked during high-visibility courtroom contests involving educational tests, their influence on members of the educational measurement community is considerable. Thus, for example, those who write textbooks about educational testing almost always try to make sure what they are recommending in those textbooks is in accord with the latest rendition of the AERA, APA, NCME *Standards*. Such authors, indeed, universally defer to the *Standards* when recommending how to play in educational testing's sandbox.

Periodically revised, the 1999 version of the *Standards* held sway for about one and a half decades, because until mid-2014 the 1999 *Standards* were essentially

the only game in town. During that 1999–2014 period, a series of extraordinarily important applications of educational testing took place (for example, the role of students' test scores in educational accountability programs such as those fostered by NCLB and ESSA). Not surprisingly, then, the 1999 *Standards* were regarded by many educators as being somewhat out of date. So when, after a lengthy revision and review process, the 2014 edition of the *Standards* was published, great interest was predictably displayed by assessment specialists. To illustrate, if pivotal concepts regarding educational assessment were altered, or even if such concepts were more clearly explicated, these alterations and these clarified explications would, in a very few years, be incorporated into the guidelines governing not only what was being professionally recommended, but also what educators ought to be learning about educational testing.

The 2014 *Standards* do not introduce any dramatic reconceptualizations of the fundamental notions of educational testing that have guided educational measurement specialists since the 1999 version of the *Standards*. However, the new edition of this influential document both clarifies and tightens the interpretation of several key concepts in educational assessment. We will consider the most salient of those clarified "tightenings" in Chapters 3 and 4 on reliability and validity. The 2014 *Standards* did, however, more clearly emphasize the importance of assessment *fairness* than had earlier revisions. Thus, in the most recent *Standards* the three chief emphases for educational assessment are validity, reliability, and fairness.

Do classroom teachers need to become knowledgeable regarding all (or even most) of what's contained in the 2014 *Standards?* Surely not! We can let the educational measurement specialists of the world fuss with adhering to and interpreting content in the 2014 edition of the *Standards*. But it is a reasonable expectation that teachers at least realize that the ground rules of educational assessment did not arrive from outer space or from a Far Eastern measurement guru. No, these nuts-and-bolts guidelines about educational testing undergo a rigorous review, rewriting, and approval process every decade or two by committees of three national organizations most concerned with such testing. What teachers need to know, however, is that if they ever find themselves embroiled in any sort of test-related controversy, there exists an authoritative collection of definite dos and don'ts. It is called the *Standards for Educational and Psychological Testing* (2014) and it is available to all.

Assessment Versus Testing

So far, you have encountered several contrasts between teaching and testing, even though, when you glance at this book's cover, you find that it's supposed to be a book about assessment. If you're alert, you've already started to wonder, What's this author trying to pull off? Am I going to learn about *testing* or am I going to learn about *assessment?* Is *assessment* simply a more fashionable word for *testing?*

In short, what's the author up to? These are reasonable questions, and you will now be supplied with a set of compelling, confidence-engendering answers.

Almost everyone knows about the kinds of tests typically encountered in school. Most of today's adults, indeed, were on the receiving end of a hoard of teacher-dispensed tests during their own days in school. There were final exams, midterm exams, end-of-unit tests, pop quizzes, and even (in the interest of gender equity, of course) "mom quizzes." All those tests had one thing in common. They represented the teacher's attempt to get a fix on how much the teacher's students had learned. More accurately, such tests were employed to determine a student's status with respect to the knowledge or skills the teacher was attempting to promote. This is an altogether praiseworthy endeavor for teachers—to find out how much students know. If teachers are reasonably sure what their students currently know, they can more accurately tailor any future instructional activities to promote what their students truly need to learn.

The sorts of tests referred to in the preceding paragraph, such as the quizzes and examinations most of us took in school, have historically been paper-and-pencil instruments. When most of us were in school, the three most common forms of tests were essay tests, multiple-choice tests, and true-false tests. Until the past decade or so, those three kinds of tests were, by far, the most prevalent sorts of tests found in classrooms.

In recent years, however, educators have been urged to broaden their conception of testing so students' status is determined via a wider variety of measuring devices—a variety extending well beyond traditional paper-and-pencil tests. It is not merely for the sake of variety that teachers have been challenged to expand their repertoire of testing techniques. Rather, thoughtful educators have recognized that a number of important kinds of student learning are not measured most appropriately by traditional paper-and-pencil tests. If, for example, a teacher wants to determine how well students can function orally in a job interview, it's pretty clear that a written true-false test simply doesn't cut it.

Moreover, the widespread use of computers, along with all sorts of digitally based communication devices, has led to a dramatic upsurge in the use of computer-based testing. Later in the book, you will learn about not only computer-based assessments, but also computer-adaptive assessment. It almost seems that every new advance in computer capabilities is soon followed by a derivative computer-rooted version of educational testing. Although most classroom teachers will not be generating computer-controlled assessments for their own classroom use (because of the sophisticated programming demands of such test building), a teacher's students will occasionally be on the receiving end of large-scale computer-controlled standardized tests—often administered under a teacher's supervision. Accordingly, teachers should become at least reasonably conversant with the nature of these increasingly prevalent assessments. As you will learn later in this book, while there are wondrous things that can be accomplished when a test developer's inventiveness hooks up with computers' capabilities, computer-governed assessments can also be misused. It is not merely

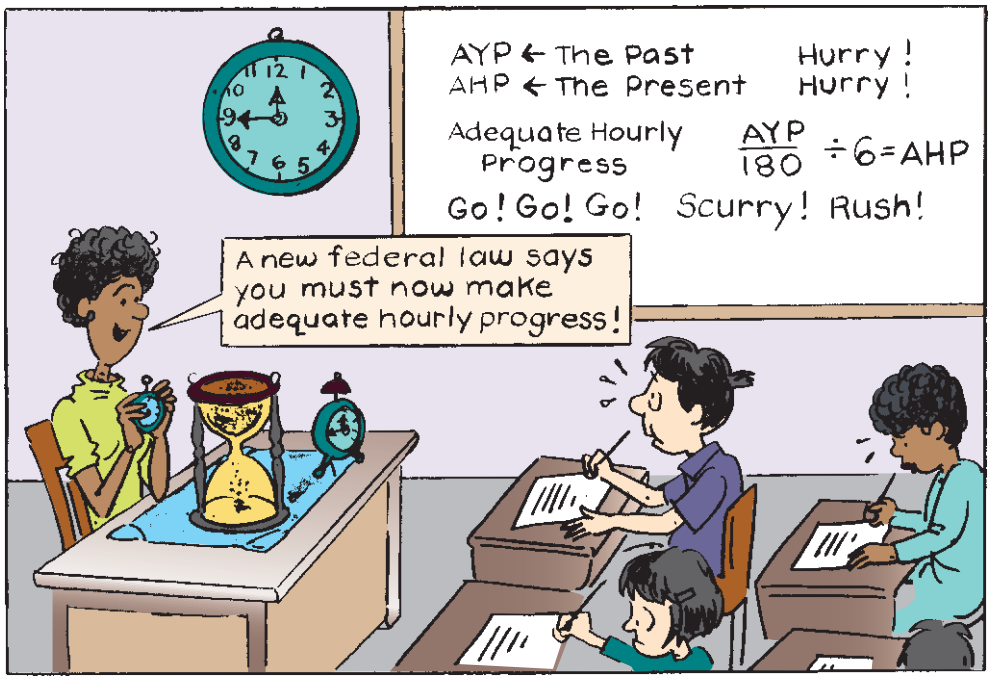
the presence of a high-powered computer that leads to appropriate assessment. Rather, assessment appropriateness depends on the specific way a computer is applied.

Thus, because many worthwhile learning outcomes are not best measured by paper-and-pencil tests, and because, when most people use the word *test* they automatically think of traditional paper-and-pencil tests, the term *assessment* has been increasingly adopted by many educators and measurement specialists. Assessment is a broader descriptor of the kinds of educational measuring teachers do—a descriptor that, while certainly including traditional paper-and-pencil tests, covers many more kinds of measurement procedures as well. Here is a working definition of ***assessment*** as it is used in an educational context:

> Educational assessment is a formal attempt to determine students' status with respect to educational variables of interest.

Lest you be put off by this fairly foreboding definition, let's briefly consider its chief elements. Note that the kind of assessment we're talking about is aimed at determining the status of students regarding "educational variables of interest." *Variables* are merely things that vary. (You probably figured this out all on your own!) In education, for example, we find that students vary in how much they know about a subject, how skilled they are in performing certain operations (such as long division), and how positive their attitudes are toward school. Those are the sorts of variables with which a teacher is typically concerned. Accordingly, they are the "variables of interest" that teachers typically measure. If the teacher's instructional focus is on the Industrial Revolution, then the teacher may wish to assess how much students know about the Industrial Revolution. In that case, the variable of interest would be the degree of students' knowledge regarding the Industrial Revolution. If the teacher is interested in how confident students are in their own written composition skills, then students' composition confidence would be a variable of interest. Educational assessment deals with such variables.

Our working definition also indicates that educational assessment constitutes a "formal" attempt to get a fix on students' status. (If you prefer, you could replace "formal" with "deliberate," "systematic," or "planned" because educational assessment is not a "seat of the pants" undertaking. Rather, it is a careful, purposive activity.) As human beings, we make all sorts of informal determinations regarding people's status. For example, we may conclude that the woman who cut into the supermarket line ahead of us is rude, or that the man who keeps stumbling as he climbs a set of stairs is clumsy. But these are informal status determinations. Teachers, too, make numerous informal judgments about their students. For instance, a teacher might conclude, on the basis of a student's glum demeanor during the first few moments of class, that the student is definitely grumpy. Such *informal* appraisals, although they may be quite useful to teachers, should not be regarded as educational assessment.



When I was a high school teacher, for example, I employed informal judgment to conclude that Raymond Gonty, one of the seniors in my U.S. government class, was not really interested in what I was teaching. I suspected this chiefly because Raymond usually slept during class. And I became more firmly convinced when he began arriving to class carrying a pillow!

The kind of educational assessment you'll be reading about in this book is formal—that is, it's a deliberate effort to determine a student's status regarding such variables as the student's knowledge, skills, or attitudes. The kind of educational assessment you'll be considering is more than a teacher's "impressions." Rather, you'll be learning about systematic ways to get an accurate fix on a student's status.

Assessment, therefore, is a broad and relatively nonrestrictive label for the kinds of testing and measuring teachers must do. It is a label to help remind educators that the measurement of students' status should include far more than paper-and-pencil instruments. Assessment embraces diverse kinds of tests and measurements. In the remaining pages, you'll find that, although the label *assessment* is used often, the terms ***test*** and ***measurement*** are sometimes used interchangeably with it. This does not represent any subtle nuances regarding the topic at hand—rather, it reflects an author simply getting tired of using the A-word.

Why Should Teachers Know About Assessment? Yesteryear's Answers

Let's indulge in a bit of time travel. Suppose you were magically transported back to the 1950s or 1960s. And, as long as we're in a let's-pretend mode, imagine you're a new teacher taking part in a fall orientation for first-year teachers in a large school district. The thematic topic of the particular session you're attending is "Why Should Teachers Know About Testing?" The session's lecturer, Professor Tess Tumm, is supplying the audience with a set of traditional answers to this thematic question based on how teachers actually can use classroom tests. Because you are a docile new teacher (remember, this is imaginary), you are compliantly taking notes to help guide you during the coming school year.

What's being suggested, as you've probably guessed, is that there are a number of fairly traditional answers to the question of why teachers should learn about assessment. Those answers have been around for several decades. There is also a set of more current answers to the question of why teachers should know about assessment. Let's give tradition its due and, initially, consider four time-honored answers to the question. Although these reasons for knowing about classroom assessment may have been around for a while, they're still compelling because they are rooted in the realities of what skilled teachers can do with classroom assessment. These four reasons may well have been the major points treated by Professor Tumm during our imaginary orientation session of yesteryear.

Determining Students' Current Status

One important reason why a teacher might assess students is to determine what they presently know and can do—for instance, what a group of students' current levels of knowledge are or what their current cognitive skills happen to be. If, for example, a teacher has been instructing students about a series of mathematical operations, there are moments during an instructional sequence when it would be useful for the teacher to know which of those operations have been mastered—and which haven't. Based on students' test performances, then, a teacher can decide which mathematical operations seem to need more instructional attention and which seem to need only a brief review.

There is one oft-encountered instance in which teachers can benefit considerably by using tests to determine students' current status, and it comes up many times during a school year. When teachers are trying to promote their students' attainment of knowledge or skills that are relatively *new* to the students, it is remarkably helpful to get a fix on what it is that students already know and can do.

For instance, if Jaime is already proficient in solving simultaneous equations, it's a waste of Jaime's time to make him plow through practice piles of such equations. In my youth, the following expression was sometimes used: "That's like carrying coal to Newcastle." It was used to disparage any scheme that reeked of redundancy. (One assumed that coal mining was a big deal in Newcastle.) Well, teachers who relentlessly keep instructing students in knowledge or skills that the students already possess are definitely carting coal to Newcastle. Assessment, if it is working the way it is supposed to work, allows teachers to identify students' current capabilities and, as a consequence, can help teachers avoid superfluous and wasteful instruction.

Thus, by measuring students' status, teachers can discern (1) where to put their instructional energies to ameliorate a student's shortcomings and (2) what already mastered skills or knowledge can be instructionally avoided. Such assessment is particularly useful for a teacher's planning if the assessment is carried out at the beginning of an instructional sequence. This kind of early diagnosis is often referred to as ***preassessment*** because it is assessment that takes place prior to the teacher's initiation of instruction.

Monitoring Students' Progress

A second, related answer to the question of why teachers should assess is that such assessments help teachers determine whether their students are making satisfactory progress. Sometimes, of course, it's easy for teachers to tell whether their students are progressing satisfactorily. I can still recall, with suitable embarrassment, the absolutely scintillating lesson I provided as a high school English teacher on the topic of modifying gerunds with possessives. It was a lesson designed for a full-class period, and I was confident that, at its conclusion, my students would not only understand the topic but also be able to explain to others why one of the following sentences contains an appropriate pronoun and one does not:

Improper
Pronoun *Gerund*
↓ ↓

☹ **Sentence 1:** I really appreciate you sending the brownies.

Proper
Pronoun *Gerund*
↓ ↓

☺ **Sentence 2:** I really appreciate your sending the brownies.

At the end of a bravura 40-minute lesson, replete with all sorts of real-life examples and a host of on-target practice activities, I was certain I had effectively taught students that because a gerund is a noun-form of a verb, any modifiers of such gerunds, including pronouns, must be possessive. And yet, at the end

of the lesson, when I looked into my students' baffled faces, I realized that my optimism was unwarranted. After asking several students to explain to me the essence of what I'd been talking about, I quickly discerned that my lesson about gerund modifiers was not an award-winning effort. Most of my students couldn't distinguish between a gerund and a geranium.

Although teachers can occasionally discern informally, as I did, that their students aren't making satisfactory progress, more often than not we find teachers' believing their students are progressing quite well. (Note in the previous sentence that the modifier of the gerund *believing* is the possessive form of *teachers*. Yes, I'm still trying.) It's only human nature for teachers to believe they're teaching well and their students are learning well. But unless teachers systematically monitor students' progress via some type of assessment, there's too much chance that teachers will improperly conclude that progress is taking place when, in fact, it isn't.

A useful function of classroom assessment, therefore, is to determine whether students are moving satisfactorily toward the instructional outcomes the teacher is seeking to promote. If progress for all students is satisfactory, of course, then the teacher need make no instructional adjustments. If progress for most students is satisfactory but a few students are falling behind, then some separate doses of remedial assistance would seem to be in order. If progress for most students is inadequate, then the teacher should substantially modify whatever instructional approach is being used. Progress monitoring is a time-honored and altogether sensible use of classroom assessment.

A teacher also ought to monitor students' progress via classroom assessment because, more often than you'd think, the teacher can stop instructing on a certain topic well in advance of what the teacher had anticipated. Suppose, for instance, you're attempting to get your students to acquire a certain skill, and you've set aside 2 weeks to promote their mastery of that skill. If you monitor students' progress with an assessment after only a week, however, and discover your students have already mastered the skill, you should simply scrap your week 2 plans and smilingly move on to the next topic.

Another way of thinking about the monitoring of student progress is that it positions teachers to use the results of classroom tests as part of *formative assessment*—that is, the use of assessment-elicited evidence intended to improve unsuccessful yet still modifiable instruction. *Summative assessment*, in contrast, refers to the use of tests to make a final success/failure decision about a relatively unmodifiable set of instructional activities. In a review of research studies focused on the instructional payoffs of formatively oriented classroom assessment, two British investigators (Black and Wiliam, 1998) concluded that the use of progress-monitoring classroom assessments can promote striking gains in student learning on both teacher-made and external exams.

Based on the Black and Wiliam conclusions about the major instructional dividends of formatively oriented classroom assessment, members of Britain's Assessment Reform Group introduced the idea of classroom assessment *for*

learning—in contrast to assessment *of*, learning. They describe this approach as follows:

> Assessment for learning is any assessment for which the first priority in its design and practice is to serve the purpose of promoting pupils' learning. It thus differs from assessment designed primarily to serve the purpose of accountability, or of ranking, or of certifying competence. (Black et al., 2002)

Stiggins and Chappuis (2016) have also pushed assessment for learning as the cornerstone of effective classroom measurement. Later, in Chapter 12, you will learn much more about the fundamentals of formative assessment.

Assigning Grades

If we were somehow able to carry out an instant nationwide survey of beginning teachers and asked them, "What is the most important function of classroom assessment?" The answer we'd get from many of the surveyed teachers is eminently predictable. They'd respond: *to give grades*.

That's certainly what I thought testing was all about when I taught in public schools. To be honest (confession, I am told, is good for the soul), the *only* reason I tested my students was to give them grades. I've talked to hundreds of teachers during the past few years, and I've been dismayed at how many of them continue to regard grade giving as testing's *dominant* function. A third reason, therefore, why teachers assess students is to assemble the evidence necessary to give their students grades. Most school systems are structured so that the end-of-course or end-of-year grades a student earns constitute the beginnings of a record of the student's personal accomplishments—a record destined to follow the student throughout life. Thus, it is imperative that teachers not assign grades capriciously. Whether we like it or not, students' grades are important.

The best way to assign grades properly is to collect evidence of a student's accomplishments so the teacher will have access to ample information before deciding whether to dish out an A, B, C, D, or F to a student. Some school systems employ less traditional grading systems—for example, the use of descriptive verbal reports that are relayed to parents. Yet, whatever the reporting system used, it is clear that the teacher's assessment activities can provide the evidence necessary to make sensible student-by-student appraisals. The more frequent and varied the evidence of student accomplishments, the more judiciously the teacher can assign to students the grades they deserve.

A corollary principle linked to "tests as grade determiners" is that some teachers also employ the prospect of upcoming tests to motivate their students. Because a student's grade is often dependent on the student's test performances, teachers frequently employ admonitions such as "Be sure to study this chapter carefully, because you have an important end-of-chapter exam coming up on Thursday!" Some teachers surely employ comments about impending tests as a motivational device.

In recent years, several thoughtful educators have proffered sensible guidance regarding how teachers ought to award grades to their students (for example, Guskey, 2015). A consensus of these writers' thinking—a consensus focused on "standards-based" grading—will be presented in Chapter 16 to wrap up this edition of *Classroom Assessment*.

Determining One's Own Instructional Effectiveness

A fourth and final reason why teachers have traditionally been told they should test students is that students' test performances can help teachers infer how effective their teaching has been. Suppose a teacher sets out to have students master a set of worthwhile skills and knowledge regarding Topic *X* during a 3-week instructional unit. Prior to instruction, a brief test indicated students knew almost nothing about Topic *X*, but after the unit was concluded, a more lengthy test revealed students had mastered most of the skills and knowledge addressed during the Topic *X* unit.

Because the comparison of students' pretest and posttest results indicated the teacher's students had acquired ample knowledge and skills regarding Topic *X*, the teacher had a convincing chunk of evidence that the instructional approach being used appears to be working. If the teacher's instruction seems to be promoting the desired outcomes, then it probably shouldn't be altered much.

On the other hand, let's say a teacher's Topic *X* pretest-to-posttest results for students suggest that students' progress has been minimal. Comparing results on the end-of-instruction posttest to students' performance on the preinstruction test reveals that students barely know more than they knew before the instruction commenced. Such trivial student growth should suggest that the teacher make adjustments in the instructional activities when teaching Topic *X* again next term or next year.

This does not imply that students' pretest-to-posttest results are the only way for teachers to tell whether they're flying or flopping instructionally, but students' end-of-instruction performances on assessment devices constitute a particularly compelling indication of whether teachers should retain, alter, or jettison their current instructional procedures.

In review, then, we've considered four fairly traditional answers to the question of why teachers should assess students. Here they are again:

Traditional Reasons Why Teachers Assess Students

- To determine students' status
- To monitor students' progress
- To assign grades to students
- To determine instructional effectiveness

You will notice that each of these four uses of educational assessment is directly related to *helping the teacher make a decision*. When a teacher assesses to determine students' status, the teacher uses test results to *decide* what instructional

objectives to pursue. When a teacher assesses students' progress, the teacher uses test results to *decide* whether certain parts of the ongoing instructional program need to be altered. When a teacher assesses students to help assign grades, the teacher uses students' performances to *decide* which students get which grades. And, finally, when a teacher uses pretest-to-posttest assessment results to indicate how effective an instructional sequence has been, the teacher is trying to *decide* whether the instructional sequence needs to be overhauled. Teachers should never assess students without a clear understanding of what decision will be informed by the results of the assessment. Indeed. the chief function of educational assessment is to improve the quality of educational decision making.

Taken in concert, the four traditional reasons just described should incline teachers to assess up a storm in their classrooms. But these days, even more reasons can be given to explain why teachers need to know about assessment.

Why Should Teachers Know About Assessment? Today's Answers

In addition to the four traditional reasons why teachers need to know about assessment, there are three new reasons that should incline teachers to dive joyfully into the assessment pool. These three reasons, having emerged during the past decade or so, provide compelling support for the conclusion that today's teachers dare not be ignorant regarding educational assessment. Let's consider three new roles for educational assessment and see why these new functions of educational testing should inspire you to pump up your assessment knowledge and skills.

Influencing Public Perceptions of Educational Effectiveness

When I was a high school teacher a long while ago, teachers were occasionally asked to give nationally standardized achievement tests. But, to be honest, no one really paid much attention to the test results. My fellow teachers glanced at the test-score reports but were rarely influenced by them. The public was essentially oblivious to the testing process and altogether disinterested in the results, unless, of course, parents received a report that their child was performing below expectations. Testing took place in the 1950s and 1960s, but it was no big deal.

During the 1970s and 1980s, however, a modest journalistic wrinkle changed all that. Newspaper editors began to publish statewide educational test results on a district-by-district and even school-by-school basis. Citizens could see how their school or district stacked up to other schools or districts in the state. Districts and schools were *ranked* from top to bottom.

From a news perspective, the publishing of test results was a genuine coup. The test scores were inexpensive to obtain, and readers were really interested in the scores. Residents of low-ranked districts could complain; residents of high-ranked

districts could crow. More important, because there are no other handy indices of educational effectiveness around, test results became the measuring stick by which citizens reached conclusions about how well their schools were doing. There are many reports of realtors trying to peddle homes to prospective buyers on the basis of their being located "in a school district with excellent test scores."

As matters stand, students' performances on a state's accountability tests are certain to influence the way that *all* teachers are evaluated—even if a particular teacher's own students never come within miles of an accountability test. Here's how this will happen.

Suppose you teach ninth-grade social studies, and your ninth-graders aren't required to take federally required accountability tests. Suppose you're a second-grade teacher, and your students aren't required to take any sort of accountability test. Suppose you're a high school teacher who teaches subjects and grade levels where no federal or state accountability tests are required. In all these "suppose" situations, *your* students won't be taking accountability exams. However, the public's perception of *your* personal effectiveness will most certainly be influenced by the scores of *your* school's students on any accountability tests that are required *and reported* for such schools. Let's be honest—do you want to be a teacher in a "failing" school? Do you want your students' parents to regard *you* as ineffective because you happen to do your teaching in what's thought to be a sub-par school?

The reality is that the performance of any school's students on federally stipulated accountability tests will reflect on every teacher in that school. If you teach in a school that's regarded as successful, then you will be seen as a member of an effective educational team. The opposite is also true. Unless federal accountability requirements are substantially softened, no public school teacher will be able to remain isolated from the impact of externally imposed accountability tests.

And, as will be pointed out later in the book, the nature of a school's success on high-stakes external assessments, such as federally required accountability tests, will (and, indeed, *should*) have an impact on the sorts of classroom assessments you personally choose to employ. Today's educators live in an era when public perceptions of schooling are more important than many educators might prefer. Yet, like it or not, that's the reality current teachers must face.

Helping Evaluate Teachers

Teaching skill is coming under increasing scrutiny these days. With the push for more rigorous evaluation of a classroom teacher's performance, we now see many teacher appraisal systems in which students' test performances constitute one key category of evidence being used to evaluate teachers. Sometimes, teachers are directed to assemble pretest and posttest data that can be used to infer how much learning by students was promoted by the teacher. And, of course, teachers whose students are *required* to take a state's annual accountability tests understand all too well that their students' scores on those tests will play a prominent role in teacher evaluation—that is, in the evaluation of *their* teaching.

Decision Time

Pressure from "Higher Ups"

Laura Lund has been teaching second-graders at Horace Mann Elementary School for the past 3 years. During that period, Laura has become increasingly convinced that "developmentally appropriate instruction" is what she wants in her classroom. Developmentally appropriate instruction takes place when the instructional activities for children are matched not only with the typical developmental level of children in that grade but also with the particular developmental level of each child. Because of her growing commitment to developmental appropriateness, and to its clear implications for individualized instruction, Laura's students now no longer receive, in unison, the same kinds of massed practice drills in reading and mathematics that Laura provided earlier in her career.

Having discovered what kinds of changes are taking place in Laura's second grade, however, the third-grade and fourth-grade teachers in her school have registered great concern over what they regard as less attention to academics, at least less attention of the traditional sort. Because state accountability tests are given to all third- and fourth-grade students each spring, Laura's colleagues are afraid their students will not perform well on those tests because they will not be skilled at the end of the second grade.

A year or so earlier, when Laura was teaching her second grade in a fairly traditional manner, it was widely recognized that most of her students went on to the third grade with a solid mastery of reading and mathematics. Now, however, the school's third- and fourth-grade teachers fear that "Horace Mann's accountability scores may plummet."

As Laura sees it, she must decide whether to (1) revert to her former instructional practices or (2) maintain her emphasis on developmentally appropriate instruction. In either case, she realizes that she must try to justify her action to her colleagues.

If you were Laura Lund, what would your decision be?

Although we will consider the topic of teacher evaluation far more thoroughly in Chapter 15, it should be noted at this point that a pair of federal initiatives have spurred much greater use of students' test scores in the appraisal of teachers. In 2009, the federal Race to the Top Program offered some serious financial grants to states that would be willing to install, among other reforms, teacher evaluation systems in which students' test performances played a prominent role. Two years later, in 2011, once again federal officials offered a flexibility program that allowed states to seek a waiver from the harsh penalties linked to the final days of the No Child Left Behind Act if they installed teacher evaluation programs in which students' test scores were regarded as a significant factor in evaluating the state's teachers.

Even though the education officials of most states signed up for one or both of those federal incentive programs and promised to implement systems for evaluating teachers (and principals) using programs featuring students' assessed growth, a good many states now seem to be treading water or scurrying away from the implementation of their once-promised educator evaluation programs. Nonetheless, in most of our states, descriptions of the current state-decreed teacher

evaluation system call for use of students' *measured* growth as one key evaluative criterion.

As a practical matter, then, because educational assessments will be employed to collect evidence of students' learning, and because this evidence will be used to evaluate teachers, a teacher would have to be a downright dunce to dodge the acquisition of information about sensible and senseless ways to measure students' status.

However, as you will learn in later chapters, only certain kinds of educational assessments can properly carry out this sort of test-based task. Most of the tests proposed for this purpose are altogether inappropriate for such an evaluative assignment. Nonetheless, if judgments about teachers' quality are—because of well-intentioned legislative actions—to be based in part on students' assessment performances, then it is apparent that teachers need to learn about the kinds of tests that will support, or possibly distort, this evaluative endeavor.

Experienced teachers will be quick to tell you that the caliber of students' test performances is dramatically influenced by the caliber of the students being tested. It should be apparent that a teacher who is blessed with a flock of bright students will almost always get better test results than a teacher who must work with a less able group of students. And let's not forget about the quality of students' previous teachers. Wouldn't you rather be receiving a new group of students who had been effectively taught by Mrs. X than a group of students who had been ineffectively taught by Mrs. Y? Nonetheless, increasing numbers of statewide and districtwide teacher evaluation systems now call for teachers to assemble tangible evidence of student accomplishments based on external exams or teacher-made classroom assessments. It is clear, therefore, that today's teachers need to know enough about educational assessment to corral compelling evidence regarding their own students' growth. In Chapter 15 we will consider today's teacher evaluation strategies in more detail.

Clarifying Teachers' Instructional Intentions

For many years, educational tests were regarded as instructional afterthoughts. As soon as an instructional unit was over, the teacher got busy cranking out a test. Tests were rarely created before instruction was initiated. Instead, tests were devised *after* instruction to fulfill some of the traditional functions of educational assessment described earlier in the chapter—for example, the assignment of grades.

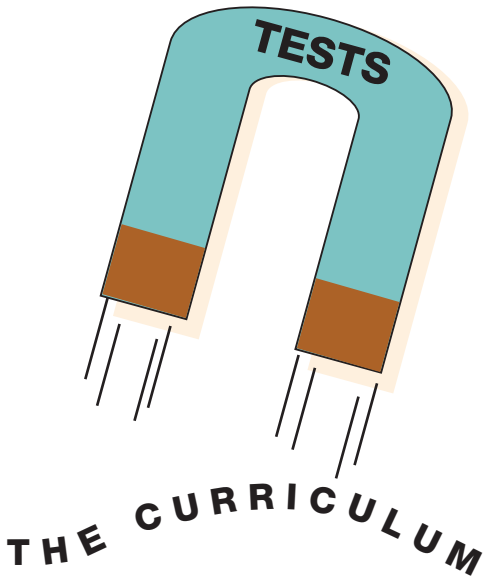
Today, however, many educational measuring instruments have become high-stakes tests. A ***high-stakes test*** is an assessment for which important consequences ride on the test's results. One example of an educational high-stakes test is a statewide test of basic skills that must be mastered before a student graduates. (Note that the important consequences are for the test taker.) Another example is the results of a districtwide achievement test that are publicized so local taxpayers' judgments about their schools' educational effectiveness are influenced by the test results. (Note that in this second case, the important consequences affect the educators who prepared the students, not the test takers themselves.)

A federally required accountability test falls into this second category of high-stakes tests. Because students' performances on these tests will so powerfully influence the way people regard a school staff's quality, such accountability tests will be genuinely high-stakes tests. You should know, however, there is nothing in ESSA that requires diploma denial or obliges students to be held back at grade level if they fail to perform well enough on a test. A *state's* decision, however, can transform a federal test into one that has an adverse impact on particular categories of students. Many people continue to be confused by this, because they assume that any federally mandated accountability test automatically requires diploma-denial or promotion-denial testing. It's just not so.

Insofar as important consequences are directly linked to assessment results, the content of such high-stakes tests tends to be emphasized instructionally by teachers. Because teachers want their students to perform well on high-stakes tests (for the students' own good and/or for the teacher's benefit), high-stakes tests tend to serve as the kind of curricular magnet seen in Figure 1.1.

On some educational grounds, teachers might prefer that tests did not influence instruction so directly, but the reality is that high-stakes assessment will almost certainly have an impact on classroom instructional practices. Because this curricular influence is sure to be present, it will be in teachers' and students' best interests if the nature of the upcoming assessment is well enough understood so that the teacher can organize the most effective, on-target instruction possible. (Later in the book, we will consider the deficits of teaching exclusively in terms of assessment targets.) In a sense, however, the more that teachers understand what the innards of a test are, the more effectively they can use this understanding to clarify what's to be sought instructionally.

Figure 1.1 The Curricular Impact of High-Stakes Tests



Even the low-stakes classroom tests routinely employed by teachers can help teachers clarify their instructional targets. Tests should obviously not be instructional afterthoughts. Rather, classroom assessment instruments should always be prepared *prior* to any instructional planning, in order for the teacher to better understand what is being sought of students and, therefore, what to incorporate in instructional activities. Assessment instruments prepared prior to instruction concretely exemplify a teacher's instructional intentions and, as a consequence, help clarify those intentions. Accordingly, clarified instructional intentions characteristically lead to more effective instructional decisions by the teacher. The better you understand where you're going, the more efficiently you can get there.

To reiterate, we've now looked at three reasons why today's teachers, unlike their counterparts of a few decades ago, need to be well informed about assessment. These reasons supplement, but do not replace, the traditional reasons why teachers assess students. Here are the three new reasons for teachers to be familiar with educational assessment:

Today's Reasons for Teachers to Know About Assessment

- Test results determine public perceptions of educational effectiveness.
- Students' assessment performances are often included as part of the teacher evaluation process.
- As clarifiers of instructional intentions, assessment devices help improve instructional quality.

These reasons are also linked to decisions. For instance, when citizens use test results to reach judgments about a school district's effectiveness, those judgments can play a major role in determining what level of taxpayer support will be provided in that district. There are also decisions on the line when students' test scores are used to evaluate teachers. Such decisions as whether the teacher should be granted tenure or receive merit-pay awards are illustrative of the kinds of decisions that can ride, at least in part, on the results of educational assessments. Finally, from the teacher's perspective, when tests serve as clarifiers of the teacher's instructional intentions, the teacher can make better decisions about how to put together instructional activities likely to help students attain the instructional outcomes represented by the assessment. With these three current roles of educational assessment, just as was true with the four more traditional roles of educational assessment, test results should contribute to educational *decisions*.

What Do Classroom Teachers Really Need to Know About Assessment?

Whether you are already a teacher or are preparing to become a teacher, you really do need to know about educational assessment. But the field of educational assessment contains reams of information. In fact, some educators devote their entire careers to

assessment. Clearly, there's more to educational assessment than you probably need to know. What, then, should *classroom teachers* know about assessment?

The title of this book—*Classroom Assessment: What Teachers Need to Know*—suggests an answer. The key word in the title, at least for purposes of this discussion, is *need*. There are oodles of fascinating things about assessment. But to help your students learn, you really don't *need* to know a host of assessment esoterica. This book about educational assessment is deliberately focused on those things that you really *must know* in order to promote your students' learning most effectively. It is flat-out silly to clutter your skull with a galaxy of nice-to-know but nonessential knowledge about educational assessment. Such nice-to-know content often crowds out the need-to-know content. There is, after all, only so much skull-space available.

As a preview, let's consider briefly what you will have learned by the time you reach the book's index. (Because no research evidence supports any book's index as a teaching tool, if you haven't learned what's needed by the time you get to this book's index, it's likely to be too late.) It may be easier for you to get a handle on what you'll be reading if you realize you'll be covering topics dealing chiefly with

1. Constructing your own assessment instruments
2. Using assessment instruments constructed by others
3. Planning instruction based on instructionally illuminating assessments

Creating Classroom Assessment Instruments

Let's start with the kinds of classroom assessment devices you will personally need to create. The chief thing you will learn in this book is how to construct a wide variety of assessment instruments you can use as part of your day-to-day classroom instruction. You really do need to know how to determine what your students have learned—for example, whether they comprehend what they have read. You also really do need to know how to get a fix on your students' educationally relevant attitudes—such as how positively disposed your students are toward the subject(s) you're teaching. Thus, you are going to be learning how to create classroom assessment approaches to measure students' *achievement* (that is, the knowledge and/or skills students acquire) as well as students' *affect* (that is, the educationally relevant attitudes, interests, and values influenced by school).

As suggested earlier, the classroom assessment procedures you'll be learning about in the following pages will extend well beyond traditional paper-and-pencil testing instruments. You may even learn about a few assessment approaches with which you are currently unfamiliar.

In a related vein, you will also learn how to judge the quality of the assessment devices *you* create. And, at the same time, you will learn how to judge the quality of assessment devices created by others. Those "others" might be your own colleagues or, perhaps, the folks who devise large-scale assessment instruments such as districtwide or statewide tests. It may seem presumptuous

to suggest that you, a classroom teacher (in practice or in preparation), could be judging the efforts of folks who create large-scale standardized educational tests. But you'll discover from this book that you will, indeed, possess the knowledge and skills necessary to distinguish between tawdry and terrific practices by those who create such large-scale educational tests. In particular, you can use your new knowledge to judge the quality of any accountability tests that may be used in your setting. Those tests are going to be so important that, if you are teaching in a state where accountability tests are educationally unsound, you definitely need to know it. What can you do if you discover that your state's high-stakes tests are inappropriate? Well, two action options come quickly to mind. For openers, you can learn enough about the shortcomings of the state's tests so that you are able to explain coherently to your students' parents why the high-stakes tests that your state has chosen are unsound. Second, and this may require some effort on your part, you can take get involved in educator organizations that are willing to help bring about the installation of more suitable educational assessments. You can do neither of these things, of course, if you have only a skimpy knowledge of what makes your state's tests tick—and how they really ought to be ticking.

It is also important for you to know enough about educational assessment so that you can assist your colleagues in evaluating an ever-increasing array of commercially developed educational tests. Educators have seen a spate of such tests developed in recent years. Some of these vendor-produced tests may be quite useful to teachers. Some of them, however, are seriously flawed—apparently cranked out merely to bring in a few bucks from desperate educators. There is no guarantee that a published test ought ever to have been published. Only an educator who possesses at least a small sack full of assessment sophistication will be able to tell whether a commercially created educational test is yummy or gummy.

Fundamentally, educational assessment rests on a foundation of common sense. Once you learn the technical vocabulary of assessment, you'll be able to identify departures from common-sense assessment practices, whether those departures are seen in your own tests, in the tests of a teacher down the hall, or in the tests created by district, state, or national assessment specialists. In short, after you finish this book, you really won't need to defer to any "measurement experts." You'll know enough to spot serious shortcomings in their efforts.

In Chapter 3, for example, you will learn about three criteria you can use to evaluate all educational tests. Those criteria apply with equal force to tests you might personally develop and to tests that are commercially developed. Once you get the hang of how to evaluate educational tests, you can apply this evaluative skill to any educational test you encounter.

Interpreting Standardized Test Results

Because your students will often be assessed with nationally developed or state-developed standardized tests, you will need to know how to interpret the results of such tests. In general, commercially developed educational tests focus either on (1) students' *achievement*, which, as noted earlier, deals with the knowledge

Assessment Literacy: Only for Grown-Ups?

Students' performances on tests can have an enormous impact not only on the students who take educational tests but also on the test taker's family. Teachers, too, are often affected by their students' test scores. To illustrate, today's teachers seem to be frequently buffeted by educational accountability tests whereon students' test scores can have an impact on teachers' tenure, assignment, and salaries. Clearly, educators at all levels, whether teachers or administrators, need to learn enough about educational tests to carry out their responsibilities successfully.

What those educators need, then, is a reasonable dose of ***assessment literacy***. And here's a definition of it:

> Assessment literacy consists of an individual's understandings of the fundamental assessment concepts and procedures deemed likely to influence educational decisions.

Notice that this definition is focused on someone's understandings of educational measurement's *basic* concepts and procedures of educational assessment. What an assessment-literate person needs to understand is not esoteric and incomprehensible. Rather, most of it is just common sense applied to educational measurement. Describing assessment literacy a bit differently, it represents the main-line measurement procedures and concepts thought to make a difference in the decisions made about the students who take educational tests.

Well, if that's what assessment literacy is, who needs to have it? There's considerable pressure these days on teachers to become more assessment literate. You are currently reading a book that, unless the book's author has really let you down, ought to help you personally become more assessment literate. But what about educational policymakers? And what about parents of school-age children? And, finally, *what about students themselves?* Don't all three of these groups need to beef up their understandings about the key assessment-related principles and processes that can influence students' lives?

Here's a little challenge for you. As you read this book, occasionally pause to think how you might relay to policymakers (such as school board members), parents (such as your own students' parents), and students (such as the ones you're teaching) the most essential things about the assessment-related concepts and procedures you're encountering. Remember, because test results these days can increasingly enhance or impair the decisions made about students, don't those students *at least* have the right to know what's going on behind the assessment curtain? A reasonable helping of assessment literacy is good for almost everyone!

It should be noted that, recently, increasing numbers of educational leaders have advocated the acquisition of ***data literacy*** by educators. Whereas assessment literacy is exclusively focused on understandings about the use of educational tests, the notion of data literacy typically also covers an educator's comprehension of a broader range of educational evidence, typically numerical in nature, regarding schools, districts, or states. To illustrate, students' attendance and tardiness are both variables of interest to educators—variables that can contribute to key decisions about educational programs—but such data are not captured by educational testing devices. Thus, the label "data literacy," because it can also include attention to scads of test results, is more inclusive than "assessment literacy." As a practical matter, however, both of those literacy genres are eminently worthy of advocacy.

and skills that students have acquired in school, or on (2) their *aptitude*, which is a term used to describe a student's academic *potential*. You should know, however, that the term *aptitude* is falling from grace these days. Several decades ago, people comfortably talked about "intelligence." Prospective teachers learned all about the *intelligence quotient* (IQ), which was a numerical way of indicating the degree to which a particular individual's intellectual abilities exceeded or fell short of conventional expectations for such individuals.

To calculate someone's IQ, you simply divided a student's *mental age* (based on how well a student's test score stacked up against a norm group's scores) by the student's *chronological age* (based on a nearby calendar). The result of this division was that student's IQ:

$$\frac{\text{MA}}{\text{CA}} = \text{IQ}$$

But "intelligence" has fallen out of favor with educators during the past few decades. The term *intelligence* conveys the notion that students possess an inborn, *intractable* potential over which schools have little influence. Yet the so-called intelligence tests, widely used until recently, often measured what students had learned at school or, more important, what students had learned at home. Thus, the term *aptitude* has been increasingly used rather than *intelligence* to convey a notion of a student's academic potential. But even the term *aptitude* tends to create the perception that there is some sort of innate cap on one's potential. Because of this perception, the commercial test makers who formerly created so-called intelligence tests, and then renamed them aptitude tests, are looking for a less negatively loaded descriptor. Interestingly, the tests themselves, although they've often be re-christened, haven't really changed all that much.

At any rate, you'll learn how to make sense out of the kinds of reports regarding student performance released by those who conduct large-scale assessments. You will need this knowledge not only to inform your own decisions about classroom instruction, but also to interpret students' test performances to parents who may demand answers to questions such as, "What does my son's standardized test performance at the 40th percentile really mean?" or "If my fifth-grade daughter earned a grade-equivalent score at the eighth-grade level on this year's standardized achievement test, why shouldn't she be promoted?" In short, you'll learn how to interpret students' performances on both ***achievement tests*** and ***aptitude tests***. Moreover, given the relatively recent arrival of computer-administered and computer-adaptive educational tests, you'll find that parents are apt to be raising questions about such technologically abetted tests. Teachers need to be able to respond to such questions—preferably with the correct answers.

Instructionally Illuminating Assessment

Earlier, we suggested that because assessment devices exemplify a teacher's instructional intentions, those assessment instruments can clarify the teacher's instructional decision making. You'll learn more about how the link between

But What Does This Have to Do with Teaching?

This chapter contains over half a dozen reasons why teachers need to learn about assessment. Actually, there are seven reasons presented in the chapter, and that's one more reason than a half-dozen. (Notice how low-level the arithmetic in this book is going to be!)

But let's single out the *two* reasons why, *from an instructional perspective*, all teachers need to know about testing. The first of these reasons is the last of the seven reasons cited in the chapter—namely, the *instructional-planning payoffs* teachers can get from a clarified understanding of what they're trying to have their students accomplish. Because a properly constructed classroom test can truly exemplify what a teacher is trying to achieve, the resulting *clarity of intention* helps teachers make more astute decisions when they plan their instruction.

The second reason why all teachers should become more astute regarding assessment is also instructionally rooted. It's the second of the four traditional reasons considered in the chapter—namely, so teachers can *monitor students' progress*. If teachers use students' assessed levels of achievement to determine whether the current instructional plan is stellar or sickly, then teachers' adjustments in lessons can, if warranted, be more accurately made. Without the evidence yielded by classroom assessments, a teacher will often fail to spot instructional inadequacies. As the British investigators Black and Wiliam quite clearly reported, the instructional dividends from monitoring students' progress can be striking. Their views were based on solid research investigations, not wishful yearnings. And of course, any assertion delivered in a British accent simply reeks of credibility.

testing and teaching can prove beneficial to your students because you will be able to provide more on-target, effective instruction.

On the other hand, you'll also learn how some teachers inappropriately prepare their students for tests, particularly for high-stakes tests. You will learn about ways of judging whether a given test preparation practice is (1) in students' best interests from an educational perspective and (2) in educators' best interests from an ethical perspective. In short, you'll learn about the increasingly important relationship between instruction and assessment.

Later, you'll discover ways to build classroom assessments so that they'll have a decisively positive impact on how well you teach. *Tests*, if deliberately created with instruction in mind, *can boost your personal success as a teacher*. We'll dig into that topic in Chapter 12.

There's another important issue that should be brought to your attention—namely, the possibility (after you've finished the book) of your helping *parents* learn more about educational assessment. And why, you might ask, should a teacher be messing around trying to promote parental measurement moxie? It's a good question. And the answer is this: Parents who are *assessment literate* will be better able to *help you help their children learn* more successfully.

You see, most parents know little more about testing than what they can recall, often vaguely, from their own classroom days as students. But the nature of classroom testing has changed dramatically since that time. Not only are new

approaches to assessment being used in classrooms, but students' test scores are also being used to judge the success of teachers' instructional efforts. You'll learn in Chapter 15 that, depending on the tests being used, this may be a dumb idea. If you and your students' parents truly understand the fundamentals of educational assessment, you can work together in many ways that will benefit your students. Assessment-literate parents can be a potent force to counter the serious misuses of educational tests we see so often today. And if you are teaching in a setting where officials have opted to use instructionally inappropriate accountability tests, you'll find that assessment-literate parents can be a potent political force that might, if you're lucky, help get more appropriate accountability tests installed.

There's another audience for assessment literacy that you should consider as you wend your way through this book. Please recognize that the lives of today's students are increasingly influenced by their performances on various kinds of educational tests. Why not, therefore, administer at least a dose of assessment literacy to students themselves? As you will see, most of the assessment concepts treated in this book are not particularly complicated. Indeed, the truly essential assessment understandings needed by students are well within the grasp of those students. Why not provide them with a bit of assessment literacy? They really need it!

Parent Talk

Mr. and Mrs. Smothers are attending a back-to-school night at a middle school where their daughter, Cathy, is a fifth-grader. After briefly leafing through Cathy's math portfolio and language arts portfolio, they get around to the real reason they've come to school. Mrs. Smothers, looking more than a little belligerent, says, "Cathy tells us she gets several teacher-made tests in class every week. All that testing can't be necessary. It obviously reduces the time you spend teaching her! Why is there so darn much testing in your class?"

If I were you, here's how I'd respond to Mrs. Smothers:

"I suppose it might seem to you that there's too much testing going on in my class, and I can understand your concern about testing time taking away from teaching time. But let me explain how the time my students spend doing classroom assessments really leads to much *better* use of instructional time.

"You see, the way I use classroom assessment is to make sure my instruction is on target and, most important, to make sure I don't waste the children's time. Last month, for instance, we started a new unit in social studies, and I gave students a short pretest to find out what they already knew. To my delight, I discovered that almost all of the students—including Cathy—knew well over half of what I had been planning to teach.

"Based on the pretest results, I was able to shorten the social studies unit substantially and spend the extra time giving students more practice on their skills in map interpretation. You probably saw some of the maps Cathy was interpreting as part of her homework assignments.

"Mr. and Mrs. Smothers, I want Cathy's time in class to be as well spent as possible. And to make sure of that, I use formal and informal classroom tests to be certain that I'm teaching her and her classmates what they really need to learn."

Now, how would you respond to Mrs. Smothers?

Chapter Summary

In this chapter, the emphasis was on why teachers really need to know about assessment. Early in the chapter, the assessment-related features of various reauthorizations of the Elementary and Secondary Education Act of 1965 were briefly described, because this oft-revised federal law's impact on most teachers' instructional and assessment decisions is becoming profound. *Educational assessment* was defined as a formal attempt to determine students' status with respect to educational variables of interest. Much of the chapter was devoted to considering why teachers must become knowledgeable regarding educational assessment. Based on teachers' classroom activities, four traditional reasons were given for why teachers assess: (1) to determine students' status, (2) to monitor students' progress, (3) to assign grades, and (4) to determine a teacher's own instructional effectiveness. Based on recent uses of educational assessment results, three more reasons why teachers need to know about assessment were identified. Those more recent functions of educational tests are (1) to influence public perceptions of educational effectiveness, (2) to help evaluate teachers, and (3) to clarify teachers' instructional intentions. Regardless of the specific application of test results, however, it was emphasized that teachers should use the results of assessments *to make better decisions*. That's really the only excuse for taking up students' time with assessment.

The chapter identified three major outcomes to be attained by those reading the book—namely, they should become more knowledgeable about (1) how to construct and evaluate their own classroom tests, (2) how to interpret the results of standardized tests, and (3) how to teach students to master what's assessed in classroom and high-stakes tests. It was also suggested that an assessment-literate teacher should attempt to promote parents' and students' assessment literacy.

MyLab Education Self-Check 1.1

MyLab Education Self-Check 1.2

MyLab Education Application Exercise 1.1: Understanding Legislation

MyLab Education Application Exercise 1.2: Persuasive Writing

References

Andrade, H. (Ed.). (Winter 2009). "A special issue on classroom assessment," *Theory into Practice*, *48*, no. 1.

Black, P. J., Harrison, C., Lee, C., Marshall, B., and Wiliam, D. (2002). Working inside the black box: Assessment for learning in the classroom. London, UK: King's College London School of Education.

Black, P., and Wiliam, D. (1998). "Inside the black box: Raising standards through classroom assessment," *Phi Delta Kappan 80*, no. 2 (October): 139–148.

Curwin, R. L. (2014). "Can assessments motivate?" *Educational Leadership*, *72*, no. 1 (September): 38–41.

Goodwin, B. (2014). "Better tests don't guarantee better instruction," *Educational Leadership*, *71*, no. 6 (March): 78–81.

Gotch, C. M., and French, B. F. (2014). "A systematic review of assessment literacy measures," *Educational Measurement: Issues and Practice*, *33*, no. 2 (Summer): 14–18.

Guskey, T. R. (2015). *Challenging the conventions of grading and reporting.* Bloomington, IN: Solution Tree.

Hess, F. M. (2008/2009). "The new stupid," *Educational Leadership*, *66*, no. 4 (December/January): 12–17.

Layton, L. (June 7, 2014). "How Bill Gates pulled off the swift Common Core revolution," *The Washington Post.*

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standards for English Language Arts and Mathematics.* Washington, DC: Author.

Nichols, S. L., and Berliner, D. C. (2008). "Testing the joy out of learning," *Educational Leadership*, *65*, no. 6 (March): 14–18.

Popham, W. J. (2009). *Unlearned lessons: Six stumbling blocks to our schools' success.* Cambridge, MA: Harvard Education Press.

Stiggins, R. J., and Chappuis, J. (2016). *An introduction to student-involved assessment FOR learning* (6th ed.). Boston: Pearson.

Wiliam, D. (2013). "Assessment: The bridge between teaching and learning," *Voices from the Middle*, *21*, no. 2 (December): 15–20.

A Testing Takeaway

Assessment Literacy: What Is It, Who Needs It?*

W. James Popham, University of California, Los Angeles

It is becoming increasingly difficult these days to take part in any conversation about public education without someone's referring to *assessment literacy*—usually as a plea in support of it. But what is "assessment literacy"? And who ought to be assessment literate?

"Literacy," through the years, has referred to the ability to read and write. More recently, the term has come also to describe knowledge or competence in a specified area—such as one's "wine literacy." It is in this sense that most educators now conceive of assessment literacy, namely, a person's knowledge or competence regarding educational testing. What it is that different people need to know about educational testing, of course, varies from situation to situation. To illustrate, here's a definition of assessment literacy intended for teachers and other educators: *Assessment literacy consists of an individual's understanding of the fundamental assessment concepts and procedure deemed likely to influence educational decisions.*

As you can see, this educator-focused conception of assessment literacy addresses the *basics* of educational testing and, specifically, those fundamental concepts and procedures that are apt to have *an impact on educational decisions*. Thus, a teacher or administrator who wants to be assessment literate need not know as much about testing as an expert who specializes in measurement. Instead, educators merely need master a handful of assessment understandings—the understandings likely to impact educational decisions.

Other non-educator groups (for instance, school board members or parents of school-age children) typically need less knowledge about educational testing than those who are making daily decisions about students, because some of those decisions depend on students' test scores.

But *why* does anyone need to be assessment literate? The answer to this key question is that without the necessary knowledge regarding educational testing, *mistakes*—sometimes serious ones—are made regarding how to educate children. Examples of such mistakes, typically made by those who are *not* assessment literate, are listed as follows:

- *An Error of Commission*: A school board decides that an *ineffective* school serving students from affluent families is *successful* based on the wrong accountability test.
- *An Error of Omission*: Teachers in a middle school fail to use the formative-assessment process in their classrooms, despite its research-ratified record of success.

Because assessment-literate educators make fewer test-based errors of commission or omission than teachers who are less well-informed, please do what you can to promote the wider acquisition of assessment literacy among all education stakeholders. Indeed, increasing assessment literacy may be the most cost-effective way to improve our schools.

Chapter 2
Deciding What to Assess

Learning Outcomes

2.1 Identify key factors impacting teachers' decision-making when determining measurement targets for classroom assessments.

2.2 Identify key components of the Common Core State Standards when determining measurement targets for classroom assessments.

Chief Chapter Outcome

An ability to identify, and understand the impact of, key factors that can help teachers determine the measurement targets for their classroom assessments

If classroom assessment does not help teachers do a better job of educating their students, then classroom assessment has no reason to exist. Consistent with that premise, this book about classroom assessment is really about teaching, not testing. That's because well-conceived classroom assessment will almost always lead to better taught students. So, before getting into some of the measurement concepts typically linked to educational testing, this chapter addresses an overridingly important *educational* question that each teacher must answer about classroom assessment. This crucial question, suitably highlighted in uppercase type, is

WHAT SHOULD A CLASSROOM TEACHER TRY TO ASSESS?

In this chapter you'll encounter a collection of factors that teachers should consider before deciding what sorts of things their classroom assessments should measure.

What to Assess

Far too many teachers simply stumble into an assessment pattern without giving serious consideration to why they're assessing what they're assessing. Many teachers, for example, test students in order to dispense grades in a manner that somehow reflects the levels of academic performances that students have displayed on those tests. Students who score well on the teacher's tests are given good grades; low-scoring students get the other kind. Traditionally, the need to

dole out grades to students has been a key factor spurring teachers to assess their students. Other teachers, of course, also employ classroom tests as motivational tools—that is, to encourage students to "study harder."

Yet, as suggested in Chapter 1, there are a number of other significant reasons why teachers construct and use assessment instruments. For instance, results of classroom assessments may be employed to identify certain students' areas of deficiency so the teacher can more effectively target additional instruction at those content or skill areas where there's the greatest need. Another important function of classroom assessment is to help teachers, understand more clearly, *before* designing an instructional sequence, what their end-of-instruction targets really are. This clarification occurs because properly constructed assessment procedures can exemplify, and thereby illuminate, the nature of instructional targets for teachers.

Decision-Driven Assessment

Teachers use tests to get information about their students. Teachers typically make score-based interpretations about their students' status with respect to whatever curricular aims are being represented by the tests. Based on these interpretations, teachers then make decisions. Sometimes the decisions are as straightforward as whether to give Steve Smith an A or a B. Sometimes the decisions are more difficult, such as how to adjust an ongoing instructional unit based on students' performances on an along-the-way exam. But whatever the decisions are, classroom assessment should be unequivocally focused on the *action options* teachers have at their disposal. The information garnered from assessing students is then used to help a teacher make the specific decision that's at issue. Because the nature of the decision to be illuminated by the test results will usually influence the kind of assessment approach the teacher selects, it is important to clarify, *prior to the creation of a test*, just what decision or decisions will be influenced by students' test performances.

It may seem silly to you, or at least somewhat unnecessary, to identify *in advance* what decisions are linked to your classroom assessments, but it really does make a difference in determining what you should assess. For instance, suppose the key decision riding on a set of test results is how to structure a series of remedial instructional activities for those students who performed poorly during a teaching unit promoting a higher-order thinking skill. In this sort of situation, the teacher would definitely need specific diagnostic information about the "building-block" subskills or enabling knowledge each student did or didn't possess. Thus, it is not sufficient for a test merely to assess students' mastery of the overall skill being taught. There would also need to be a sufficient number of items assessing students' mastery of enabling knowledge or each building-block subskill. Based on the diagnostic data derived from the test, a sensible set of remedial instructional activities could then be designed for the low performers.

If, on the other hand, the decision linked to test results is a simple determination of whether the teacher's instruction was sufficiently effective, a pretest–posttest assessment of more global (and less diagnostic) outcomes would suffice. It's also appropriate, when judging the quality of a teacher's instruction, to assess students' *attitudes* relevant to what was being taught. We'll look into how you can measure your students' attitudes in Chapter 10.

In short, the decisions to be informed by assessment results should always influence the nature of the assessments themselves. Teachers should, therefore, routinely consider the decision(s) at issue prior to creating a classroom assessment device.

A fairly easy way to decide whether a test's results will really influence a classroom teacher's decision is to imagine that the test results turn out in two opposite ways—for example, a set of excellent student test performances versus a set of disappointing student test performances. If the teacher would be likely to make a *different* decision based on those disparate sets of performances, then the teacher truly has a test that can help inform decisions. If the teacher's decision would be pretty much the same no matter what the test results were, there's a strong likelihood that the assessment procedure will be more ritual than genuine help for the teacher.

The Role of Curricular Aims

Much classroom assessment takes place after an instructional sequence has been completed. There are exceptions, of course, such as the during-instruction monitoring of students' progress as part of the formative-assessment process (about which you'll learn lots more in Chapter 12). There's also the pretesting a teacher might use as part of a pretest–posttest evaluation of the teacher's instructional effectiveness. Yet many classroom tests are used at the close of instruction. And, because much classroom instruction is intended to help students achieve specified outcomes, it is quite reasonable to think that if, early on, teachers consider the outcomes they want their students to achieve, those teachers can more readily answer the what-to-assess question. What teachers should assess will, in most instances, stem directly from the intended consequences of instruction, because those hoped-for consequences will influence what the teacher will be teaching and, most likely, what the students will be learning.

Let's pause for just a moment to consider how best to label a teacher's instructional intentions. This might seem unnecessary to you because, after all, why should the way teachers label their instructional aspirations make any real difference? Aren't a teacher's instructional aspirations simply what a teacher hopes will happen to kids as a consequence of the teacher's instruction?

In recent years, the label "content standards" has emerged as the most fashionable descriptor for what in one locale might be called "learning outcomes" and, in another setting, "instructional goals." If you accept that reality as you rumble through this book, you might prefer to mentally slide in whatever

synonym suits you. Regrettably, there is confusion among educators about how to describe the outcomes we hope students will achieve. If you are aware of this potential confusion, you can adroitly dodge it and move on to more important concerns. If you don't understand the nature of curriculum-label confusion, however, you are likely either to become confused yourself or, worse, add to the confusion.

Getting down to basics, ***curriculum*** consists of the sought-for *ends* of instruction—for example, changes in students' knowledge or skills that a teacher hopes students will experience as a result of what was taught. In contrast, ***instruction*** is the *means* teachers employ to promote students' achievement of the curricular ends being sought. Instruction, therefore, consists of the activities a teacher has students carry out in an attempt to accomplish one or more intended curricular outcomes. Simply put, *curriculum equals ends* and *instruction equals means*. Confusion arises when educators try to plaster diverse sorts of labels on the curricular ends they hope their students will accomplish.

No widespread consensus among educators exists regarding what to call the intended outcomes of instruction. During the past half-century, for example, we've seen curricular outcomes referred to as goals, objectives, outcomes, expectations, benchmarks, and content standards. In most instances, certain labels were supposed to apply to broader, more general curricular outcomes, while other labels were given to more specific, less general outcomes. What's important when carving out what you want your classroom assessments to measure is that you must understand precisely the meaning of whatever curricular labels are being used by colleagues in your school, district, or state. In most instances, whatever the curricular labels are, they are presented (often with examples) in some sort of official curricular document. Be sure you have a handle on what sorts of curricular descriptors are being used in your own corner of the world. You'll dodge much curricular chaos if you do. Most crucially, do *not* assume that you and your coworkers are using curricular labels in the same way. More often than not, you won't be. So, in passing, just do a dandy definition dance to see whether you are all dancing to the same tune.

More often than not in this book, the label ***curricular aim*** will be employed to characterize the desired outcomes of instruction. That phrase *curricular aim* has not been around for all that long and, therefore, has not had most of its meaning leeched from it. You'll sometimes encounter in the book such phrases as "educational objectives" or "instructional objectives" to describe curricular aims—just for a descriptive change of pace. But, as indicated, you need to find out what sorts of curricular labels are being used in the setting where you teach. Odds are that those labels will be different from the descriptors used here. That's okay, so long as you know what's going on—and can clarify terminology tangles when they transpire.

Consideration of your own curricular aims, then, is one action you can take to help get a fix on what you should assess. The more clearly you state those aims, the more useful they will be to you in answering the what-to-assess question.