

---

# Introduction

**S**ince the passage of No Child Left Behind legislation in 2001 (NCLB), we have worked with teams of teachers and school administrators, discussing test items, negotiating student performance level descriptors, and, ultimately, setting **cut scores** for **accountability** purposes with personnel from state departments of education. We both have also been lucky enough to work with a state department of education as it develops a comprehensive system of accountability that includes standards for teaching, learning, and classroom formative **assessment**.

As we go about this work, we are troubled by what the teachers in the front lines tell us about how students are being prepared for test taking in this era of accountability. Many educators report feeling compelled to abandon what they know to be the best ways to teach reading in exchange for a test-preparation curriculum designed to raise test scores. Other schools have hired independent consulting firms, staffed by well-meaning, smart people who nevertheless have no classroom experience or educational background, to coach veteran teachers on how to teach test-taking strategies to increase reading scores. Why are educators so ready to turn over their professional voices and expertise?

The atmosphere the NCLB created in classrooms across the United States can explain this sudden lack of confidence. Many teachers find themselves judged by the test scores of their students—test scores that are affected by factors beyond their control: students' academic history, students' abilities, school facilities and equipment, transience of the population, socioeconomic class, and so on. It is not surprising that teachers who are threatened with pay for performance incentives or reorganization based on the student body's **adequate yearly progress (AYP)** would feel compelled to change how they teach for the promise of improved test scores, even if it means giving up strategies for teaching reading and writing that research and experience show are effective. What can teachers do in the classroom to help students prepare for a high-stakes, standardized, multiple-choice reading test without sacrificing what they know to be best practice?

This book is designed to give reading teachers practical tools to improve reading test scores without forgetting everything we've learned about the best ways to teach kids how to be literate and how to *love* reading. This is particularly important in light of the pressures to raise student test scores in reading in this age of accountability.

Our intention with this book is threefold:

1. to provide teaching tips to use in an elementary reading classroom,
2. to make sure that *all* students are prepared for the high-stakes reading tests, without causing more work for teachers, and
3. to accomplish this preparation without teaching to the test.

For some readers who have worked in elementary schools a long time, the policy background or the teaching techniques may not be new. We expect these readers to find this book a useful tool to use in professional development inservices, particularly Chapter 6. For educators charged with purchasing technology to support reading initiatives in schools, we believe Chapter 3 will be particularly helpful to you. For elementary principals who are looking for ethical strategies to raise building reading test scores without asking teachers to use skill-and-drill-worksheets, this book is for you!

NCLB has caused a myriad of unintended consequences for literacy educators, especially in schools in danger of not meeting AYP. An overemphasis on teaching the portions of the reading domain that are testable with large-scale assessments, like spelling and finding facts from a reading passage, narrows the language arts curriculum in ways that go against what teachers know is the best way to teach. For example, in order to target reading instruction at students at varying developmental levels, some schools have adopted reading curricula that limit the choices for children's literature if the books are not on a reading level that has been determined by a publishing company to be academically appropriate. Adult readers don't choose books that way, so why should we train our children that books must be challenging to be worth reading? Students don't develop a lifelong love of reading without practicing reading what interests them.

As educators, we must hold our ground as professionals and refuse to compromise our excellent tried and true teaching practices in the name of higher test scores. The tests themselves are not the problem; in fact, the test scores can provide teachers with valuable information about student performance on basic language arts skills. But, teaching only the skills that can be tested leads to a narrowing of the curriculum and to the deprofessionalization of teachers. It is crucial that the expertise of teachers be brought to bear on curriculum decisions at the classroom level.

## WHAT WE'VE LEARNED ABOUT THE TESTING INDUSTRY

Liz and Hilleary met in 2002 in an educational measurement class at the University of Iowa. We both had graduate assistantships with Iowa Testing Programs researching and developing tests for the *Iowa Tests of Basic Skills* (ITBS). Amazingly, our first year at this new job was the year NCLB was passed.

What we know is that good readers should be able to use a number of strategies to approach a novel text. These strategies can be taught in school. Test preparation worksheets are boring, and they don't make kids better readers. In fact, really good reading tests are virtually impervious to test prep. That is, if a test taker is successful on the test, it is because he or she is a good reader, not because he or she is a good test taker.

In Ketter and Pool's (2001) article, *Exploring the impact of a high-stakes direct writing assessment in two high school classrooms*, the authors write that "critics of standardized multiple choice testing believe that teaching to such tests narrows the range of activities in which students engage" (p. 344). In many school districts, the stakes placed on test scores have created an academic environment where what is tested can affect what is taught, an idea that runs counter to best practice for **psychometricians** (test developers) and teachers alike.

It may surprise some teachers to hear this, but the testing industry as a whole does not claim that student learning can be assessed only with standardized tests. For example the *Iowa Tests Interpretive Guide for Teachers and Counselors* says "many of the common misuses [of the tests] stem from depending on the scores from a single test or test battery to make an important decision about a student or a class of students," (Hoover et al., 2003, p. 11). Certainly standardized, multiple-choice tests can be used to assess reading comprehension, spelling, and grammar, but when high stakes are placed on the tests, the consequences can be profound. It forces English teachers to spend class time focusing on a slice of what teachers want their students to be able to do, because the assessment of an English teacher's teaching and of student performance is based on the standardized tests of reading comprehension and language skills. The end result can be a crowding out of all the kinds of learning that English teachers believe are important in exchange for more reading comprehension, throwing the balance of the curriculum off.

Nothing can replace teacher observation, and a single test score should never be used to make an important decision about a student or a class of students, but there are some things that standardized tests *can* do extremely well. A good achievement battery can provide teachers achievement data

on students' general reading comprehension skills that make it possible to monitor year-to-year developmental changes and can provide a basis for reports to parents that will enable home and school to work together in the students' best interests.

The standardized delivery of the assessment ensures fairness, and machine-scorable tests are an efficient and relatively inexpensive way to get a great deal of valuable information about student achievement in a short amount of time. Historically, however, when the stakes are high for teachers and students, there is a temptation for tests to drive the curriculum.

### How Standardized Tests Are Made

One way to improve communication between teachers and test makers is to make visible the test development process. Many graduates of teacher education programs enter the classroom with limited understanding of test development for classroom use or of measurement concepts like norms and grade level equivalencies. When some light is shed on how standardized tests are made, perhaps teachers and curriculum experts can discover where there are opportunities to participate in the standardized test development in their districts.

Writing a standardized test requires some training to do it well. To build a reading test, test developers look for high quality, diverse literature, on a broad range of topics, written for children and young adults that has probably not been seen by the students before in order to ensure a level playing field for all test takers. The people engaged in this search are teachers, professors, graduate students working on degrees in educational measurement, and editors. We frequent garage sales and used book stores looking for books or articles on interesting topics, and we keep a massive database of every children's book that has won a major book award or has been excerpted in a textbook to avoid choosing a text that might privilege a group of students with prior experience with that text.

You should know that some test companies "*home brew*" their reading passages. What this means is that rather than find a piece of published text, they hire writers to create an original passage on a topic of their choice. For the test developers, that makes it easy to ensure that there is enough *meat* to the text to ask questions, and it guarantees that the topic will meet all of the requirements for the test specifications. For example, if the state requires 500-word nonfiction biographical sketch at eighth grade, a test developer can save time reading biographies written for eighth graders and culling the material to 500 interesting words by just writing something original. What is more, the topic can be easily state specific (think: Abraham Lincoln in Illinois or John Glenn in Ohio).

When a customized test for a particular state is built, the content standards are used as guides to write items that align with the curriculum. When we build a test battery for academic achievement, like the *Iowa Tests of Basic Skills* (Hoover, Dunbar, & Frisbie, 2001) and *Iowa Tests of Educational Development* (Forsyth, Ansley, Feldt, & Alnot, 2001), we write more general reading comprehension items that measure a student's ability to comprehend facts in the text, to make inferences, and to generalize from what they have read, skills that separate proficient readers from developing readers.

After we write and edit items (a process that frankly takes months of wordsmithing and involves dozens of item writers and editors), we field-test them, both statewide and nationally, at multiple grade levels in order to decide which items are best suited at each grade.

The data from those tryouts are analyzed item-by-item for **bias** on race, ethnicity, socioeconomic class, sex, geographical region, and private, Catholic, or public schools. Items where bias is detected are thrown out, and items that span a range of difficulty are selected. We also analyze the item response data for optimal functioning to see what percentage of students who scored high on the test as a whole got each item correct and to be sure that each item is functioning as we intend.

Once the reading tests have been piloted and the final items have been selected, we administer the tests around the country. The data from those administrations are used to build national norms so that teachers, parents, students, and other stakeholders can get a sense of how their students are reading compared to other children like them across the country. Our norms are available on the same categories that we use in our bias review: race, ethnicity, socioeconomic class, sex, geographical region, and private, Catholic, or public schools. So, for example, a school in a rural part of the country can have norms that show how their fourth-grade students performed on the ITBS reading test compared to all of the other students in rural schools nationally.

## **So What Good Are the Data From Standardized Tests?**

The other important score that comes from the norming process is grade equivalents. When a seventh-grade student's ITBS Reading Comprehension score is compared with all of the other seventh graders who took the test in the fall of seventh grade, we can provide a report that says overall, your seventh grader performed on the test like most seventh graders in their third month of middle school. Now alone, this information probably isn't particularly interesting. But when a high school teacher looks at a student's cumulative school record and sees that every year this student scores at grade level, that teacher has some evidence of measured yearly

growth and a picture of academic progress in reading over time begins to form. This information is by no means the only data that should be presented to parents wanting to know how their child's reading comprehension is progressing. However, norms and grade equivalents are a relatively inexpensive, time-efficient, and fair way of comparing student performance nationally, not just locally.

In short, the primary reason for using a standardized achievement battery is to gather information that can be used to help improve instruction. Standardized tests do not measure all the worthwhile outcomes of an English curriculum; the diversity of instructional methods and materials makes it impractical for any test to attempt to do that. However, there are a number of generally held outcomes toward which all students are expected to progress as they go through school, regardless of the specific courses they take or the curriculum they may be following. These skills, which cut across the curriculum and may be the province of more than one department, are among the major goals of literacy education.

## CHAPTER OVERVIEWS

This book offers an approach to test preparation that does not require that teachers sacrifice everything they know about the best ways to teach kids to read. Test preparation worksheets and drill-and-kill activities do not make children into lifelong readers. Throughout the book, I provide research from the academic community to support the strategies and theories that are offered. In addition, I include stories from the field about the ways the accountability movement is corrupting teaching practice and what can be done about it.

*Chapter 1* describes ways to be sure that what you are teaching in your classroom matches what is going to be tested. Alignment studies are an important way to ensure that you and your students are getting a more accurate picture from the test scores of what kids know and are able to do. This is not the same as teaching to the test, which is an incredibly important distinction.

*Chapter 2* offers suggestions for the use of formative assessments in literacy contexts in order to check learning as you go. For the most part, any assessment can be used in a formative way, so don't be suckered in by merchants trying to sell you a magic Formative Test.

*Chapter 3* is a conversation about teaching literacy with special populations. There is a helpful table of assistive technology devices that can be used in the service of teaching literacy to students with disabilities and with English language learners. In addition, a breakdown of the differences

between accommodations and modifications is given, along with the history of policy and legislation that have guided best practice in special education in the United States.

*Chapter 4* provides a discussion of motivation research to give teachers and administrators insight into why bribery (for example, pizza parties) doesn't have the kind of impact on test scores that one might expect. Some of the research I have conducted into ways schools try to motivate children on high-stakes tests is summarized in this chapter as well.

*Chapter 5* gives an overview of ways we can connect reading units to real-world contexts to model for students what lifelong readers do. This chapter also provides suggestions for approaching test passages as a genre study in your classroom, including a conversation about how standardized, multiple-choice tests are developed.

*Chapter 6* reviews best practice in teaching reading at the elementary level, including a conversation about theory and practice, and what to do when you feel that your expertise in teaching literacy is threatened by the reading programs being implemented in your district.

*Chapter 7* concludes with implications for your teaching and with some advice for school leaders who are in a position to make a real difference with the policies they support.

## END-OF-CHAPTER QUIZZES FOR DISCUSSION

*A short vignette, followed by a quiz, appears at the end of each chapter to provide fertile ground for lively faculty discussions or book club-like conversations with your colleagues. We envision a book study group using these as the talking points to start their discussions at the beginnings of the meetings.*

*For some readers, the scenarios may seem wildly fictitious. Let me assure you that every one of the examples we give have come from practitioners who have described to me personally what has been happening in their schools. For those of you for whom the examples seem all too familiar, you have our sympathies and our support.*

## GLOSSARY

Be sure to look at the end of the book for a glossary of assessment terms. Hopefully it will serve as a resource for you as you read.