This powerful collection from the groundbreaking Rethinking Schools magazine takes high-stakes standardized tests to task. Despite overwhelming evidence that the tests are invalid ways to measure teaching and learning—and continuing signs of their unjust effects on students and teachers—“reformers” and policy makers continue to force high-stakes tests into the public schools. Through articles that provide thoughtful and emotional critiques from the frontlines of education, Pencils Down deconstructs the damage that standardized tests wreak on our education system and the human beings that populate it. Better yet, it offers visionary forms of assessment that are not only more authentic, but also more democratic, fair, and accurate.

No questions asked, Pencils Down is one of the most comprehensive, scholarly (and yet eminently readable) take-downs of the nation’s testing craze ever produced. Combining a just-right mix of quantitative analysis with the compelling personal narratives of educators trying to reach children in ways a test can never measure, this volume is a must read for teachers, administrators, parents, and policy makers.

—Tim Wise, author, White Like Me: Reflections on Race from a Privileged Son and Dear White America: Letter to a New Minority

Finally we have a book that demystifies and challenges the madness of the “almighty” testing craze! At a time when the most vulnerable students are further victimized and made to feel even more marginal, Wayne Au and Melissa Bollow Tempel have compiled a document that should be read by anyone who thinks there is something fundamentally wrong with the current emphasis on testing over quality education.

—David Stovall, associate professor, Educational Policy Studies and African American Studies, University of Illinois at Chicago

Here, in one place, is a collection that looks at standardized testing and its alternatives in exactly the ways useful to teachers, parents, and teachers of teachers! What an enormous contribution to us all. It can change what you do at your school—tomorrow.

—Debbie Meier, author of The Power of Their Ideas and Save Our Schools steering committee member

Wayne Au and Melissa Bollow Tempel bring to us a complete exposure of the damages of high-stakes testing and sensible suggestions of alternatives with compelling stories, insightful analyses, and clear action plans. This book is a wake-up call, a must read for everyone concerned about our children’s education.

—Yong Zhao, presidential chair and associate dean, College of Education, University of Oregon, and author of Catching Up or Leading the Way: American Education in the Age of Globalization

A Rethinking Schools Publication

Rethinking high-stakes testing and accountability in public schools

Edited by Wayne Au and Melissa Bollow Tempel

Illustrations by Randall Enos | Graphic Design by The Flynstitute
Pencils Down

RETHINKING high-stakes testing and accountability in public schools

EDITED BY WAYNE AU AND MELISSA BOLLOW TEMPEL
Pencils Down: Rethinking high-stakes testing and accountability in public schools
Edited by Wayne Au and Melissa Bollow Tempel

A Rethinking Schools Publication
Rethinking Schools Ltd. is a nonprofit educational publisher of books, booklets, and a quarterly magazine on school reform, with a focus on issues of equity and social justice. To request additional copies of this book and/or a catalog of other publications, or to subscribe to Rethinking Schools magazine, contact:

Rethinking Schools
1001 E Keefe Avenue
Milwaukee, WI 53212
800-669-4192
www.rethinkingschools.org

© 2012 Rethinking Schools Ltd.
First edition

Production Editor: Catherine Capellaro
Cover and book design: The Flynstitute
Cover and chapter head illustrations: Randall Enos
Proofreading: Amalia Oulahan
Indexing: Marilyn Flaig
Business Manager: Mike Trokan

All rights reserved. Except as permitted below, no part of this book may be reproduced in any form or by any means, including electronic, without the express prior written permission of the publisher, except for brief quotation in an acknowledged review. Individuals may photocopy excerpts from this book for use in educational settings (physical settings only; internet use is not permitted), as long as such material is furnished free to students or workshop participants. For sale of any portion of this book as part of course packets, contact the Copyright Clearance Center for permissions and appropriate assessment of fees. If you have any questions, contact Rethinking Schools at the above address.

Library of Congress Cataloging-in-Publication Data

Pencils down : rethinking high-stakes testing and accountability in public schools / edited by Wayne Au and Melissa Bollow Tempel.—First edition.
   pages cm.— (A Rethinking Schools publication)
A collection of articles from Rethinking Schools magazine.
Includes bibliographical references and index.
LB3051.P46 2012
371.26—dc23
2012010474
We dedicate this book to Maya Lu and Masami for bravely being the first in their school to opt out of computerized testing; and Makoto, Haden, Nebuen, and Ollie in hopes that the testing will never reach them.
Acknowledgements

We thank Rethinking Schools editors Linda Christensen, Kelley Dawson Salas, Stan Karp, David Levine, Larry Miller, Bob Peterson, Stephanie Walters, Terry Burant, Helen Gym, Rita Tenorio, Dyan Watson, Kathy Williams, Kathy Xiong, Bill Bigelow, and Jody Sokolower for their ongoing work and support in shaping this book; Catherine Capellaro for helping shepherd this book through the process; Patrick J.B. Flynn for his artistic direction and patience; Randall Enos for his creative expression; Amalia Oulahan for her attention to detail; Kris Collett, Tegan Dowling, and Mike Trokan for doing so many things that make Rethinking Schools run; and, of course, the contributors to Pencils Down for everything they are doing to challenge high-stakes testing. We’d also like to thank our partners, Mira and Gene, for understanding the importance of getting this book published.

—Wayne Au and Melissa Bollow Tempel
# Table of Contents

**Acknowledgements** ....................................................... iv  
**Introduction** ............................................................... 1  
*It’s time to put the pencils down*  
WAYNE AU AND MELISSA BOLLOW TEMPEL

## Part 1:  
**Testing, Testing, 1, 2, 3…**

**Failing Our Kids** ........................................................... 9  
*Why the testing craze won’t fix our schools*  
KATHY WILLIAMS AND BARBARA MINER

**Common Questions about Standardized Tests** ................ 12  
KATHY WILLIAMS

**Standards: Decoy or Quality Control?** ................................. 16  
ASA HILLIARD

**A Child Is Not a Test Score** ............................................... 21  
*Assessment is a civil rights issue*  
MONTY NEILL

**The Straitjacket of Standardized Tests** ............................... 31  
*Where is the standardized test that can measure passion for learning, respect for others, and human empathy?*  
TOM McKENNA

**Racism in the History of Testing: Legacies for Today** .......... 34  
ALAN STOSKOPF

## Part 2:  
**Testing Kids**

**Testing Kindergarten** .................................................. 43  
*Young children produce lots of data*  
KELLY McMAHON

**Testing Lang** ............................................................... 46  
AMY GUTOWSKI
All Work and No Play .................................................. 47
How educational reforms are hurting our preschoolers
SHARNA OLFMAN

A Dark Cloud on the U.S. Horizon .................................. 51
A teacher’s experience in England is a cautionary tale
MELISSA SCHIEBLE

Testing Our Sanity ...................................................... 55
A 4th-grade bilingual teacher shares her story
of preparing students for mandated tests
KELLY DAWSON SALAS

Edwina Left Behind ................................................... 59
SÖREN WUERTH

The Scripted Prescription .............................................. 62
A cure for childhood
PETER CAMPBELL

Testing Our Limits ...................................................... 66
MELISSA BOLLOW TEMPEL

**Part 3:**

Testing Teaching

Teaching in Dystopia .................................................. 73
Testing has a stranglehold on education
WAYNE AU

Fighting for Electives .................................................. 78
Lessons in courage
MELISSA BOLLOW TEMPEL

Standards and Testing Attack Multiculturalism ............... 83
BILL BIGELOW

Think Less Benchmarks ................................................. 90
A flawed test does more harm than good
AMY GUTOWSKI
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>About Those Tests I Gave You</td>
<td>94</td>
</tr>
<tr>
<td><em>An open letter to my students</em></td>
<td></td>
</tr>
<tr>
<td>RUTH ANN DANDREA</td>
<td></td>
</tr>
<tr>
<td>A Teacher Pushed to the Edge</td>
<td>96</td>
</tr>
<tr>
<td>SARAH KNOPP</td>
<td></td>
</tr>
<tr>
<td>Playing Smart</td>
<td>101</td>
</tr>
<tr>
<td><em>Curricular activism and resisting the script of high-stakes testing</em></td>
<td></td>
</tr>
<tr>
<td>WAYNE AU</td>
<td></td>
</tr>
<tr>
<td>Support That Can’t Support</td>
<td>106</td>
</tr>
<tr>
<td><em>One teacher’s induction program experience</em></td>
<td></td>
</tr>
<tr>
<td>ELAINE ENGEL KESWICK</td>
<td></td>
</tr>
<tr>
<td>Coming Soon to Your Favorite Credential Program:</td>
<td>110</td>
</tr>
<tr>
<td>National Exit Exams</td>
<td></td>
</tr>
<tr>
<td>ANN BERLAK</td>
<td></td>
</tr>
<tr>
<td>Responses to “Coming Soon to Your Favorite Credential Program”</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PART 4:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Testing the Tests</strong></td>
<td></td>
</tr>
<tr>
<td>Neither Fair nor Accurate</td>
<td>127</td>
</tr>
<tr>
<td><em>Research-based reasons why high-stakes tests should not be used to evaluate teachers</em></td>
<td></td>
</tr>
<tr>
<td>WAYNE AU</td>
<td></td>
</tr>
<tr>
<td>A Test Scorer’s Lament</td>
<td>132</td>
</tr>
<tr>
<td><em>Scenes from the mad, mad, mad world of testing companies</em></td>
<td></td>
</tr>
<tr>
<td>TODD FARLEY</td>
<td></td>
</tr>
<tr>
<td>The Loneliness of the Long-Distance Test Scorer</td>
<td>135</td>
</tr>
<tr>
<td>DAN DiMAGGIO</td>
<td></td>
</tr>
<tr>
<td>Fuzzy Math</td>
<td>145</td>
</tr>
<tr>
<td>MEREDITH JACKS</td>
<td></td>
</tr>
</tbody>
</table>
Apologies to Sandra Cisneros ........................................ 148
How ETS’ computer-based writing assessment misses the mark
MAJA WILSON

On White Preference ....................................................... 156
JAY ROSNER

Testing What Matters Least ......................................... 158
What we learned when we took the Praxis Reading Specialist Test
MAIKA YEIGH, ANDIE CUNNINGHAM, AND RUTH SHAGOURY

PART 5:
Resisting and Responding to High-Stakes Testing

Responding to Test-Driven Reform .............................. 167
BARBARA MINER

High-Stakes Harm ......................................................... 169
Teaching students to read tests
LINDA CHRISTENSEN

An Untold Story of Resistance ...................................... 175
African American educators and IQ testing in the 1920s and 1930s
ALAN STOSKOPF

Teaching Is Not Testing ................................................. 183
A community-led struggle to find an alternative
to California’s graduation exam
TINY (AKA LISA) GRAY-GARCIA

None of the Above ......................................................... 189
Defiant teachers show they have had enough of high-stakes testing
AMALIA OULAHAN

Taking a Stand for Learning ........................................... 194
Chicago teachers speak out against a “really bad test”
GREG MICHIE
Testing, Tracking, and Toeing the Line
A role play on the origins of the modern high school
BILL BIGELOW

Opting Out and Speaking Up
A national movement grows to opt out of state tests
PEGGY ROBERTSON

Template for Opting Students Out of State Tests

PART 6:
Beyond High-Stakes Standardized Testing

Hallmarks of Good Assessment
THE EDITORS OF RETHINKING SCHOOLS

Alternatives to Standardized Tests
BOB PETERSON AND MONTY NEILL

A Better System for Evaluating Schools and Students
FAIRTEST: NATIONAL CENTER FOR FAIR & OPEN TESTING

Multiple Measures
It’s not easy to escape the high-stakes trap,
but here are some examples of alternatives
MONTY NEILL

Their Report Card—and Ours.
How do we know if schools are doing a good job?
And how can we make them better?
PORTLAND AREA RETHINKING SCHOOLS

Basketball and Portfolios
LINDA CHRISTENSEN

Another Path Is Possible
Despite the pressures of high-stakes testing, two Chicago principals keep an eye on what matters
GREG MICHIE
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steady Work: Finland builds a strong teaching and learning system</td>
<td>255</td>
</tr>
<tr>
<td>Teaching in Dangerous Times: In this era of demands for teacher quality, it is crucial to develop culturally relevant ways to assess teachers</td>
<td>263</td>
</tr>
<tr>
<td>Teachers Teaching Teachers: In Portland, teachers work together to create teacher-centered professional development</td>
<td>271</td>
</tr>
<tr>
<td>Taking Teacher Quality Seriously: A collaborative approach to teacher evaluation</td>
<td>278</td>
</tr>
<tr>
<td>Author Bios</td>
<td>285</td>
</tr>
<tr>
<td>Index</td>
<td>289</td>
</tr>
</tbody>
</table>
High-Stakes Harm

Teaching students to read tests

LINDA CHRISTENSEN

Tests today are high stakes. Based on these numbers, students are retained, placed in summer school and remedial classes; schools are reconstituted, and teachers’ and principals’ salaries rise and fall. Students, especially those who “fail” the tests, internalize the failure, and question their ability and their intelligence. They learn to blame themselves, and some come to believe they will not succeed because they are not capable enough. As my daughter said after receiving a 3 (not competent) on an oral Spanish test, “Maybe I don’t have what it takes.” The test took away the experience of animated conversations with her host family in Cuenavaca as well as her ability to navigate Mexico City. Instead of questioning the validity of the measurement tool as an “authentic assessment,” especially compared to her experience in Mexico, she questioned herself.

So while critical teachers might stand back and say we don’t want to have anything to do with tests, you better believe that we can’t just go on with business as usual. The question for anyone who cares about kids is how do we retain our critical stance on assessments while preparing students for them? Can we “teach the tests” without compromising what we know to be true about teaching and learning?

My friends at low-achieving elementary schools have been counseled to acclimate students to tests by redesigning their regular curriculum so that students can get accustomed to multiple-choice questions. But in an activist classroom, that’s not easy. How can a role play about an important historical or social issue be reformatted into a multiple-choice activity? How does an a, b, c, d answer format encourage students to imagine themselves an interned Japanese American or a Cherokee Indian facing government-ordered removal? Teachers are also asked to mimic the more “authentic” assessments in fairly inauthentic ways. A kindergarten colleague was asked by the 3rd-grade teachers to prepare students for the state six-trait analysis scoring guide by giving them 1-6 scores on everything from lining up at the door to tying their shoes and counting.

Clearly, this isn’t the kind of teaching we want to happen in our classrooms. To achieve real gains in student knowledge and skill, we must continue to give students a rich curriculum with varied opportunities to use their learning in real world activities. This material will generate growth that may or may not be reflected in test scores, but will increase the likelihood of students seeing themselves as readers, writers, historians, scientists, mathematicians, and thinkers.

However, I live in a state that has filled our classrooms with tests—multiple choice and work samples. As a teacher and mother who has patched up the wounds test scores left behind and as the victim of a school that was reconstituted in part due to low test scores, I am a firm advocate in fighting against the over-assessment
The question for anyone who cares about kids is how do we retain our critical stance on assessments while preparing students for them? of students. But I also believe we must seize the opening to help our students critically analyze these exams and the motives that put them into practice, as well as teach them how to improve their skills and scores. Thus, we need to teach students to critique the tests as well as coach them on how to increase their performance.

Talk Back: Questioning the Assumptions of the Test

A social justice curriculum equips students to question what is often taken for granted. Tests have become as much a part of the curriculum as books. (In fact, these days there seems to be more money for testing, test preparation, and test scoring than for the books we need to teach. A good question might be: “Why are we spending so much money on testing when we need books?”) In critical classrooms, we can make testing the object of our curiosity.

Begin by questioning the origins and purpose of these tests. I ask students: Who made them? What are the tests supposed to measure? What information will be generated from the scores? Who will have access to that information and how will it be used? For example, in Oregon, students take multiple tests. One might think that these tests will be used to help teachers more accurately assess their students’ abilities or progress to improve instruction. But because the tests are given during the school year—from February to April—and the scores aren’t returned to the schools until after the school year has ended, one has to question the legitimacy of that claim.

If we really want to know how well students are doing in school, why can’t we ask teachers, students, or parents? Why don’t we look at student work? Portfolios? Dave Hamilton, an award-winning science teacher, wrote the following in the Oregonian:

[T]ests given by classroom teachers are almost never high stakes. Teachers typically use an average of test scores (never a single test score) as only one indicator of student achievement. They also rely on written work such as papers and other assignments, class participation, and special projects. Compared to the state’s single test score, teachers have a wealth of information on which to assess student achievement … [T]eachers are the only people with the intimate knowledge of student achievement over time needed to make such judgments.

Another way to scrutinize the tests is to find statistics generated from these exams. In my school district, the local newspaper happily prints scores and rates schools, but our research and evaluation department also has broken the scores down by gender and race. Have students examine the statistics. Which school in the district usually receives high scores? Which ones don’t? Is there a pattern? Are the scores related to parents’ income? Race?

This one is tricky because I don’t want to leave the students with the idea that
race and income are indicators of intelligence or the only factors determining academic achievement. It is important to examine the questions to see how the content might favor one race or one gender or one income bracket. FairTest Examiner is a good resource for this information on the SATs. They reported how students of color scored higher, for example, on passages that feature multicultural readings. [See also Jay Rosner’s article, “On White Preference,” in this book.]

Sometimes you can find these selections in your own city or state assessments. Ruthann Hartley, a former colleague from Jefferson High School, was furious after she administered the state reading test. According to Hartley, a disproportionate number of questions examined a passage and chart from *Consumer Reports* on frequent flyer benefits. Hartley noted, “This is a problematic item for teenagers, but especially for low-income students who don’t travel. Passages like this raise the question of what is being tested. If students answer incorrectly, is it because they can’t read or because they don’t have the background knowledge?”

Students might also question how the test results are used. Who benefits if they get high scores? Are students placed in honors or remedial classes? Given scholarships? Special programs? Are teachers’ or principals’ salaries tied to the results? Have students interview school and district administrators and department chairs about how students are placed in honors classes or gifted and talented programs. Are test scores the only criterion?

Asking students to become investigators prior to exam time can help put the tests in a social context, but more than that, it diminishes the size of their opponent. Students see behind the Wizard of Oz curtain and realize that no geniuses are laboring to construct these tests. The test-item creators are not people who have scrutinized the great order of the universe and then created these 20, 40, or 60 reading and math problems to determine which students will succeed in life and which ones will fail. [See also Dan DiMaggio’s chapter, “The Loneliness of the Long Distance Test Scorer,” included in this book.]

**Investigating the Origins of the SAT**

While my junior and senior students weren’t saddled with the Certificate of Initial Mastery (CIM) reading, writing, and math tests—which Oregon 3rd, 5th, 8th, and 10th graders currently take—they were having their behinds kicked by the SATs. After their encounters with these grueling tests, they fumed at me and their math teachers. “Those tests might as well have been written in Greek!” Shameka said after the exams ruined her Saturday. For my students, an investigation of the history of the SATs was as critical as teaching them how to improve their scores. The SAT/ACT scores become a brand of shame students carry long after they bubble in the last answer. If they score low on the test, they doubt themselves and wonder if they are as capable as the kid who scored higher.
To help students understand the origins of the exam and help them put the score in perspective, my class reads a chapter from David Owen’s book *None of the Above* called “The Cult of Mental Measurement.” In this essay, Owens describes the racist past of the SATs and also points out how race continues to be a factor in these kinds of standardized tests today. Students are outraged by their discoveries. (For example, the founder of the SATs, Carl Campbell Brigham, published in the same eugenics journal that Adolf Hitler wrote for.) But even without this gem of a chapter to use, getting students to investigate the origin and use of tests in their school district or state is a good place to start. (See Bill Bigelow’s “Testing, Tracking, and Toeing the Line: A Role Play on the Origins of the Modern High School” in this volume to help students develop a critique of the historical motivation behind the testing industry.)

**Examining the Tests**

Once students have gained a critical edge on the tests, we can help them improve their performance by examining both the content and the format of the tests themselves. The more they know about how the questions are put together as well as the vocabulary of the material, the better prepared they are to meet the challenge.

In my senior English class, students demystified the SATs and used their knowledge to teach others about their discoveries. We started by analyzing each of the verbal sections of the SATs. We examined the instructions, language, and “objectives” of each section. We took apart the analogies and figured out the kinds of relationships they paired. (We used *The Princeton Review: Cracking the SAT* to help us wade through and prepare. David Owens wrote the foreword.) We looked at how the language and culture of the SATs reflected the world of upper-class society with words like heirloom, inheritance, conservatory, and regatta.

After examining each section and taking the tests a few times, I asked students to construct their own tests using the culture, content, and vocabulary of our school—from sports to dance to awards. Pairs of students worked together developing questions that the entire class examined; then we put together the Jefferson Achievement Tests (JAT).

**JAT**

*Each question below consists of a related pair of words or phrases, followed by four lettered pairs of words or phrases. Select the lettered pair that best expresses a relationship similar to that expressed in the original pair.*

1. Tony: Play ::
   a) Broadway : Annie
   b) Oscar : Tom Hanks
   c) Brandon : soccer
   d) Howard Cherry : sports
After examining each section and taking the tests a few times, I asked students to construct their own tests using the culture, content, and vocabulary of our school.

2. New Growth : perm ::
   a) press : straight
   b) weave : long
   c) corn row : braid
   d) nails : fill

(The correct answer is d. When you get “new growth,” it is time for a perm. In the same way, if you wear acrylic nails and your nails grow out, you need to get a fill.)

4. Red Beans and Rice : play ::
   a) corn and tortillas : run
   b) song : dance
   c) mozzarella : cheese
   d) sonata : musical

(The correct answer is c. Red Beans and Rice is the name of a play; mozzarella is the name of a kind of cheese.)

5. Dancebelt : boxers ::
   a) shoes : socks
   b) student : teacher
   c) leotard : leg warmers
   d) prison : freedom

(Although this one changes form from concrete nouns to abstract nouns, I like the humor and use of Jefferson’s dance legacy in the answer—d. As a male dancer explained, a dancebelt secures a male dancer’s privates during dance. Comparing the dancebelt to boxers is like comparing prison to freedom.)

After completing the test, students took the JAT up to Ruth Shagoury’s education classes at Lewis & Clark College. My students asked the pre-service teachers to imagine that the JAT was a high-stakes test that will determine their future—what college they get into, scholarships, etc. After the tests, students discussed the issues of testing and language. In this way, my students had a real audience whose future teaching practice was hopefully enlightened by their work.

Obviously, our JAT is not an exact equivalent of the SAT, but developing their own analogies from Jefferson’s culture helped students understand the mechanics of the exam. It also made them see that if they were the test makers, using their culture and their vocabulary, they could also devise a test that could be used to exclude some and include others.

Perhaps the most important lesson of the unit came when students asked,
“Why would someone want to devise a test to keep students out of a college they want to enter?” Indeed.

Teaching students to examine the history and motives of local and state tests and preparing them for the big day(s) is no substitute for fighting to end the encroachment of assessments in our classrooms. Although the work I’ve proposed may raise student scores by a few points and help students question the tests’ legitimacy as well as their results, our bigger work as teachers and parents is to engage in the battle to stop testing that makes young people, like my daughter, question their ability.