

Introduction

Since the early 1990s, one word has come up more often than any other in discussions about education. One word has filled the halls of schools, Congress, and state legislatures. One word has changed how schools work in ways perhaps more profound than any other in the nation's history. That word is *accountability*.

The policies that fall under the accountability banner—especially high-stakes testing—have been both a rallying cry and the target of condemnation. Educators, while generally accepting the need for accountability, have real and justifiable concerns about the details of particular policies. They want what is best for their students but believe that accountability has some, mostly unintended, negative consequences. Accountability policies like No Child Left Behind (NCLB) affect their professional lives as teachers and reduce their autonomy and opportunities for creativity.

Policy makers and the general public, on the other hand, paint with broader strokes and see accountability as a matter of more versus less. They see education and teachers' jobs as relatively simple. Many view the teaching profession from a business perspective, where accountability, as it is typically conceived, is more common. While acknowledging that businesses differ from schools, they still expect to see businesslike policies and practices. They might sympathize with educators' concerns, but their sympathy gets overshadowed by a powerful desire to see a "bottom line" and demonstrable results.

In business, accountability often emphasizes the individual worker. Following this logic, President Obama and state and local leaders have taken the NCLB focus on *school* performance a step further, to individual teachers and principals. In this view, if a teacher is not generating high test scores, the teacher needs to go. If a school is persistently failing, the principal needs to go. But educators express concern that schooling is more than the sum of the individual teachers and principals, and point in turn to interdependencies in the business world—among companies' marketing, production, and engineering departments, for example—to prove their point.

It is little wonder that accountability has created so much conflict. When educators express concerns, accountability advocates hear intransigence. When advocates express support for simplistic solutions, educators hear ignorance. In this world of sound bites, none of this is very surprising, but the result is deeply problematic. By paying insufficient attention to the basic nature of education, policy makers have unwittingly put in place an accountability system that is counterproductive for students and teachers alike, and often runs counter to the deeply held beliefs and goals policy makers themselves have for education.

Value-added measures have become a key topic of contention in these accountability debates. What are these measures? The short answer is that such measures are estimated contributions to student test scores made by educators. This type of measure has gained notoriety of late because the value-added approach, in theory, provides an accurate estimate of what each *individual teacher* contributes to student learning.

Nowhere have value-added measures received greater notoriety than in Los Angeles. In August, 2010, the *Los Angeles Times* announced it had obtained student achievement data from the Los Angeles Unified School District and commissioned a researcher to estimate value-added for thousands of the district's teachers. This type of analysis was not especially new, but what was new—and disturbing—was that the newspaper published the individual teachers' value-added scores along with their names. This was an unprecedented move. Professional athletes, realtors, and a handful of other professions make individual performance measures publicly available, but for teachers—and most businesspeople—disclosing this information doesn't make much sense in general. Why? Most obviously, these measures capture only teachers' contributions to student standardized test scores, which are at best loosely related to other important outcomes like creativity and engagement. Also, there are considerable errors in value-added measures. Finally, even if the value-added reports had been accurate, placing teacher performance measures on public websites will do more to wreak havoc than help students. Value-added measures grade teachers on a bell curve, so that no matter how good the entire pool of teachers is, someone will always be at the bottom and half, by definition, will always be below average. Thus half the parents looking at the scores of their children's teachers are bound to be disappointed—and many will complain

to the administration. How exactly does this help students? The *Los Angeles Times* had not thought this through beforehand or was more concerned about making headlines than improving schools.¹ Either way, this is exactly why educators distrust the ability of policy makers to design appropriate accountability policies and of the media to accurately portray school performance.

My intention in this book is to find a more productive middle ground, one that uses value-added measures as one part of a system of performance measures and accountability that improves not only test scores, but teaching and learning. Over the past several years, I have been invited to give dozens of presentations on value-added measures, most recently by educators in Los Angeles in the immediate aftermath of the unfortunate *Los Angeles Times* website release. I have found that, as the temperature has risen on value-added and accountability, so has the confusion and mistrust. The need for clarity and productive solutions will only continue to grow.

The first step is answering a basic question: How exactly are value-added measures created? Answering this question is one of my main goals for the first chapters of this book. The fact that a comprehensive explanation of value-added measures requires several chapters itself suggests that there is some genuine complexity involved. But the confusion has also been brought on by the absence of reports or books that explain the intuition and basic calculations and the practical issues that arise when using value-added measures for accountability. Instead, what we have is an academic library of hypertechnical reports that discuss value-added in excruciating statistical detail and, at the same time, ignore the many practical realities educators and policy makers face in interpreting and using the measures. Most of the topics discussed here have been discussed elsewhere, but nowhere have they been brought together in a single place in a way that is integrated, clear, and accessible to educators and policy makers.² My aim is to fill that void. If policy makers are going to build value-added measures (or really use student test scores at all) to evaluate teachers and schools, then they need to understand what they are dealing with, and educators need to know how they are being evaluated. I think educators today feel as though they bought a new cell phone and realized that the manufacturer sent not a user's manual, but the engineering specifications.

This book is a user's manual for value-added. I provide a balanced picture of the strengths and weaknesses of these measures. I provide concrete recommendations based on evidence and principles, and address misconceptions. I also explain how the stakes attached to a measure should be proportional to the quality of the measure. What is a quality measure? We cannot answer that question without considering one basic rule.

THE CARDINAL RULE OF ACCOUNTABILITY

The cardinal rule of accountability is to *hold people accountable for what they can control*. This might seem like common sense, but I will show that current accountability systems—particularly test-based accountability systems—generally violate this rule. By contrast, value-added performance measures are designed to hold educators accountable only for what they do control and contribute to student learning.

RULE 1—The cardinal rule of accountability

Hold people accountable for what they can control.

So, what do teachers control? The answer is that there is very little over which they have complete control, and those items vary by school, district, and state. Teachers control what they do to prepare for class. They control their lesson plans and the feedback they provide to students. They also control how they manage their classrooms and respond to students' individual needs.

While they have some control over what takes place in the classroom, what happens there is ultimately a negotiation between the teacher and the fifteen to forty or so other people in the room—the students. Teachers lead, prod, nudge, encourage, and occasionally punish—all in the hopes that students will respond and learn. Sometimes it works. Sometimes not. Classrooms are complicated places, and every one of them is different. All teachers have their own story about the horrible year they had because “little Johnny” was so disruptive that there was little time left for instruction—or the wonderful year when the class really “got it” for reasons unknown.

We often forget the central role in education played by the students themselves. Students' receptiveness to learning opportunities is affected

not only by their individual temperaments, but by their peers and home and community factors. For this reason, as I show below, students' knowledge and skills at any given point in time are substantially determined by factors that are largely *outside* the control of educators.

Class sizes, assistance from colleagues, school and district leadership, funding for textbooks and supplies, and community support are also largely outside teachers' control. Increasingly, teachers are losing control over the curriculum as scripted curricula become more common under high-stakes accountability. Teachers control most of their own actions, but they do not control the actions of others.

Likewise, school principals have some control over how they spend their time but limited control over what happens in classrooms and in students' lives. Good principals lead, inspire, and organize the work of the school. But once the classroom door is closed, teachers maintain considerable autonomy. Principals control many important decisions, such as who gets hired and whether to suspend students. But they have little control over budgetary matters, the quality of the building, and, in many cases, even the staff they direct. We could say the same thing about control by superintendents, district staff, and just about anyone else associated with the school system.

What this means is that many people, including the children themselves, control outcomes and this makes it difficult to measure performance in a way that follows the cardinal rule. How do we hold an individual teacher accountable for students when home and community factors have such a large impact on student learning? How do we hold a student's current school accountable when the student has attended other schools—in some cases, what seems like a parade of other schools?

Yet despite the challenges, there is growing agreement that some form of accountability is necessary and that performance measures can and should play a role. Even among teachers, a group that has generally been the most skeptical about test-based accountability policies, 76 percent believed in 2009 that making it easier to dismiss ineffective teachers would improve teacher effectiveness, and 32 percent believed that tying rewards such as salary to measured performance would do the same.³ While this is far from universal support, the fact that even this many teachers think some form of accountability might work is somewhat surprising, given

the acrimony of the public debate. Younger teachers, who represent the future of the profession, are even more optimistic about teacher accountability systems.⁴ There is clearly an opening here to hold people accountable in ways that educators can embrace.

The question is, how do we design accountability systems in smart ways? Do we use student test scores to measure school and teacher performance? If so, how? And what exactly do we do with those measures? In designing accountability policies, how do we deal with imperfections in performance measures and the fact that some important student outcomes cannot be measured well? While educators seem open to the idea of using performance measures for accountability, the difficulties faced in answering these questions lead many to support the accountability concept while opposing many of the specific accountability policies that get proposed and adopted.

The devil is in the details, and this book addresses some of the key details. In particular, I address one fundamental problem that has been largely under the radar: the problem with test-based accountability is not only the significant limitations of the tests themselves, but that the tests do not accurately represent the contributions teachers and schools make to student achievement. As a result, we are unnecessarily misjudging school and teacher performance, which is not only unfair to educators but to the students who bear the brunt of the harm.

Value-added measures offer a potential way out of this accountability dilemma. While we can use individual student test scores to diagnose student needs, value-added allows us to go further and evaluate how well teachers and schools are addressing those needs.

PURPOSE AND PREVIEW OF WHAT'S TO COME

The purpose of this book, as the title suggests, is to clear away the fog to explain what value-added measures are and how they might be used productively to improve teaching and learning.⁵ It is a user's manual, not an engineering document.

I will return to the cardinal rule of accountability again and again. The rule has two parts. The first—"Hold people accountable . . ."—acknowledges that accountability *is* important. The second—" . . . for what they

can control”—means that we need to avoid the age-old critical flaw of holding educators accountable for factors outside their control.

But following the cardinal rule is more difficult than it sounds. To see why, we have to start from the beginning. In chapter 1, I place the issue of educator performance and accountability in context by discussing some examples from outside the field and summarizing the evidence and arguments that underlie the current drive for value-added measures.

In chapters 2 through 4, I expand on the short definition of value-added measures to provide a more detailed picture. Chapter 2, describes in more detail how we currently mismeasure student achievement and then misuse these measures for accountability. I use real achievement data from the “Oakville” school district to help bring to life some of the key points of this and subsequent chapters.

In chapter 3, I discuss the primary ways of measuring student growth and propose a paradigm shift away from the traditional approach of looking at “growth” across different groups of students toward focusing on the growth of *individual* students as they progress through school. Finally, in chapter 4, I cover the details of creating value-added measures chapter 4. I define basic value-added as simple comparisons of average student growth between similar individual schools. Advanced value-added involves accounting for many factors outside the control of schools and relies on statistical predictions about what student achievement would have been if students had instead gone to the typical school. I highlight these approaches using figures and some numerical examples from Oakville.

These first four chapters collectively explain the potential value of value-added performance measures in their ideal form. Chapters 5 through 8 concentrate on the challenges that arise in applying value-added measures with real data and in real accountability settings. Chapter 5 focuses on the general statistical issues surrounding value-added, especially the distinction between systematic and random errors as well as type I and type II errors. We cannot properly interpret value-added measures without understanding these various ways in which the measures, and the educational decisions based on them, can go wrong.

In chapter 6, I explain the rationale for improving performance measures for individual teachers and the role that value-added might play in that process. Discussion of teacher value-added measures is critical because

the federal government, as well as some state governments and school districts, are pursuing value-added as the basis for teacher merit pay and tenure decisions. The research about value-added measures discussed in chapter 7 provides a mixed story about the accuracy of value-added measures, reinforcing perceptions that, compared with snapshot measures, value-added measures have less systematic error but more random error. In chapter 8, I discuss the policy debate surrounding value-added, particularly how arguments against value-added-based accountability have been rooted in a researcher perspective that is really not appropriate for policy decisions.

The last three chapters focus on solutions to some of these challenges and clearing away some of the misconceptions that prevent the effective use of value-added measures. In chapter 9, I show how the criticisms can be partly addressed through a set of recommendations about how value-added measures should be used to improve overall performance measurement. How we plan to use value-added measures also determines how we create and report them; this is the focus of chapter 10.

As I have seen in my presentations and workshops, there are common misconceptions about value-added that will hinder their use if not addressed (see “Misconceptions About Value Added”). Therefore I conclude, in chapter 11, by explaining where these misconceptions come from. Though involving some elements of truth, they are not the most useful ways to look at value-added measures.

While I aim to objectively describe value-added and its limitations and possibilities, let me explicitly acknowledge two of my opinions, both of which I believe are widely supported and backed up by evidence. First, as I have already stated, I think we can improve the way we hold educators accountable. Moreover, I think it is possible to do so in a way that most key stakeholders—including educators—would approve of. There is simply a great deal of self-inflicted mistrust on all sides of these education and accountability debates that has made improvements difficult to accomplish.

Second, I believe value-added measures of performance can improve teaching and learning—not just increase test scores, but improve the practice of teaching and encourage the genuine learning that the vast majority of parents, students, policy makers, and educators want to see. This does

Misconceptions About Value-Added

1. We cannot evaluate educators based on value-added because:
 - a. Different teachers have different students.
 - b. Value-added measures have flaws.
 - c. Student tests are inadequate.
 - d. Teaching is complicated.
 - e. Student needs are diverse.
2. Value-added is fair for teachers but not for students.
3. Value-added measures are not useful because they are summative rather than formative.
4. Because value-added measures involve comparing teachers to one another, and there are no absolute value-added standards, they are not useful for accountability.
5. Because we know so little about the effects of value-added, we cannot risk our kids' futures by experimenting with it.
6. Value-added is too complicated for educators to understand.
7. Value-added simply represents another step in the process of "industrializing" education, making it more traditional and less progressive.
8. Value-added is a magic bullet that by itself will transform education.

not mean I advocate, say, using *teacher* value-added measures for merit pay. In reality, I think the jury is still out on that particular idea. But accountability isn't just about compensation, nor is it necessarily about a focus on individual teachers. I have long advocated accountability based on *school* value-added measures. There are many reasons to think that replacing school-level snapshots with school value-added measures would produce noticeably better accountability. But it is difficult to say for sure because we are just beginning to examine the alternatives.

The temperature on value-added measures is rising because they are increasingly being used, under various names and forms, in many school

districts across the country. And as greater awareness of these accountability issues has developed in policy-making circles, it is increasingly likely that value-added measures will soon be required by the federal government for evaluating whole schools. If these changes are to have any positive impact on schools, the policy makers in charge of designing accountability need to understand the tool they are working with and educators, as the subjects of that accountability, need to understand the meaning of the measures intended to capture their performance and guide their careers. Nobody will respond well to performance-based accountability if it is neither trusted nor understood. My goal is to improve understanding of value-added and, in the process, improve the design of educational accountability systems.

Excerpt from *Value-Added Measures in Education*
by Douglas N. Harris (Harvard Education Press,
2011), pp 1-10.

This material has been reprinted with permission
of the Harvard Education Press for personal use
only. Any other use, print or electronic, will
require written permission from the Press.

For more information, please visit
www.harvardeducationpress.org.

Copyright © 2011 by the President and Fellows
of Harvard College. All rights reserved.