

# In Defense of Educators

The Problem of Idea Quality, Not “Teacher Quality”



BY E. D. HIRSCH, JR.

People who emphasize teaching quality and the central importance of teachers are right to do so. Where some go wrong is in thinking that teacher quality is an innate characteristic. The effectiveness of a teacher is not some inherent competence, as the phrase *teacher quality* suggests. Teacher effectiveness is contextual. I have witnessed over and over that in a coherent school most teachers can become highly effective.\*

Why has the topic of teacher quality suddenly reached such a crescendo? Education reform has been on the national

agenda since 1983, the year of *A Nation at Risk*. Only in the last few years has the teacher quality issue risen to the top. I think it may be reform fatigue, possibly desperation. We are blaming teachers because of our disappointments with the results of our reforms.

## A History of Misguided Reforms

The “back-to-basics” and “whole-school reform” strategies disappointed. The state standards movement and the No Child Left Behind law have left high school students just about as far behind as they were before the law was instituted. Charter schools, despite their laudable triumphs, are highly uneven in quality.<sup>1</sup> Their overall results are not much better than those of regular schools.<sup>2</sup> When favored educational ideas do not pan out

*E. D. Hirsch, Jr., is a professor emeritus at the University of Virginia and the author of many articles and books, including the bestsellers Cultural Literacy and The Schools We Need. He is a fellow of the American Academy of Arts and Sciences and the founder of the Core Knowledge Foundation. Excerpted with permission from E. D. Hirsch, Jr., Why Knowledge Matters: Rescuing Our Children from Failed Educational Theories (Harvard Education Press, 2016), [www.hepg.org/hep-home/books/why-knowledge-matters](http://www.hepg.org/hep-home/books/why-knowledge-matters).*

\*My defense of teachers does not extend to nonperforming ones. Children and the community come first. Most teachers agree. As American Federation of Teachers President Randi Weingarten has said: “If someone can’t teach after being prepared and supported, he or she shouldn’t be in our profession.”

as hoped, reformers understandably think: “The flaw is not in my theory; it must lie in poor implementation (i.e., it must be the fault of the teachers).”

But the most likely cause of disappointing results from the various reforms is that they have been primarily structural in character. They have not systematically grappled with the grade-by-grade specifics and coherence of the elementary school curriculum. Educational success is defined by what students learn—the received curriculum. Not to focus on the particulars of the very thing itself has been an evasion that is not of the teachers’ doing. The underlying theory of the reforms (reflected in state reading standards) has been that schools are teaching skills that can be developed by any suitable content. That mistaken theory has allowed the problem of grade-by-grade content to be evaded. It was that fundamental mistake about skills that has allowed teachers to be blamed for fundamental failures—the failures of guiding ideas, not of teachers.

Elementary school teachers are people who for the most part love children, who want to devote their lives to children’s education, but many find themselves stymied and frustrated in the classroom. They apply the notions received in their training, and do what they are told to do by their administrators, under the ever-present threat of reading tests that do not actually test the content that is being taught. Under these extremely unfavorable conditions of work, it’s no wonder that teacher unions have pushed back. When the classroom, which should be a daily reward, becomes a purgatory, one turns to contract stipulations.

It’s true that in the United States, there has been a deep problem with teacher preparation for more than half a century. We have a system that, according to teachers themselves, does not prepare them adequately for classroom management or the substance of what they must teach.<sup>3</sup> Therefore, my counterthesis to the blame-the-teachers theme is blame the ideas—and improve them.

The “quality” of a teacher is not a permanent given. Within the American primary school, where curriculum is neither coherent nor cumulative, it is impossible for a superb teacher to be as effective as a merely average teacher is in Japan, where the elementary school content is coherent and cumulative. For one thing, the American teacher has to deal with big discrepancies in student academic preparation, while the Japanese teacher does not. In a system with a specific and coherent curriculum, the work of each teacher builds on the work of teachers who came before. The three Cs—cooperation, coherence, and cumulativeness—yield a bigger boost than the most brilliant efforts of teachers working individually against the odds within a topic-incoherent system. A more coherent system makes teachers better individually and hugely better collectively.

American teachers (along with their students) are, in short, the tragic victims of inadequate theories. They are being blamed for intellectual failings that permeate the system within which they must work. The real problem is idea quality, not teacher quality. The difficulty lies not with the inherent abilities of teachers but with the theories that have watered down their training and created an intellectually chaotic school environment based on developmentalism, individualism, and the skills delusion. The complaint that teachers do not know their subject matter would change almost overnight with a more specific curriculum

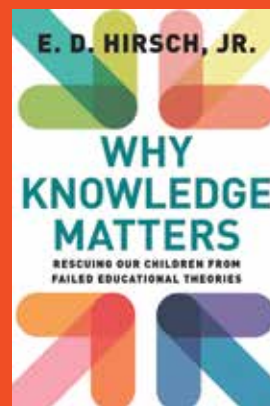
and with less evasion about what the subject matter of the curriculum ought to be. Then teachers could prepare themselves more effectively, and teacher training could ensure that teacher candidates have mastered the content they will be responsible for teaching.

A focus on technological solutions alone is also inadequate. Those who hope to find amelioration of the “teacher quality problem” through the use of computers and “blended learning” may be fostering yet another skills delusion. Such fixes haven’t worked in the past. Computers seem to work best in helping older students learn specific routines. No doubt, well-thought-out computer programs can help teachers do their work, especially for teachers in their first years. But there are *inherent*

## American teachers are being blamed for intellectual failings that permeate the system within which they must work.

limitations. For example, after decades of work and billions spent, computers cannot accurately translate from one language to another. Probably they can’t even in theory.<sup>4</sup>

Such current limitations do not lend confidence that they can transform primary education. Young students rely on an empathetic personal connection that not even our most advanced computer-adaptive programs can deliver. This is not to say that computers have no important place; it is to say that their place is supplemental, not transformative. They need to be used in support of teachers under a coherent cumulative curriculum. Computers cannot magically replace the hard thinking and political courage needed to create one.



*Why Knowledge Matters*, by E. D. Hirsch, Jr., is published by Harvard Education Press, which is offering *American Educator* readers a 20 percent discount on the purchase of the book through February 15, 2017. To order, visit [www.harvardeducationpress.org](http://www.harvardeducationpress.org) or call 888-437-1437 and use sales code WKMAFT.



## The Problem with Value-Added Teacher Evaluation

In the face of unfair scapegoating, teachers have understandably become demoralized by being constantly blamed for failures not their own. Here is the new conventional wisdom about teachers taken from the nonpartisan policy magazine *Governing* of June 13, 2013:

The research is clear: Teacher quality affects student learning more than any other school-based variable (issues such as income and parental education levels are external). And the impact of student achievement on economic competitiveness is equally clear. That's why it's so disturbing that in 2010, the SAT scores of students intending to pursue undergraduate education degrees ranked 25th out

of 29 majors generally associated with four-year degree programs. The test scores of students seeking to enter graduate education programs are similarly low and, on average, undergraduate education majors score even lower than the graduate education applicant pool as a whole. Education schools long have accepted under-qualified students, then offered them programs heavy on pedagogy and child development and light on subject-matter content.

$$A_{g,i} = \theta A_{g-1,i} + \tau_j + S\phi + X\gamma + \varepsilon$$

where  $A_{g,i}$  is the achievement of student  $i$  in grade  $g$  (the subscript  $i$  is suppressed throughout);  $A_{g-1}$  is the prior year student achievement in grade  $g-1$ ;  $S$  is a vector of school and peer factors;  $X$  is a vector of family and neighborhood inputs;  $\theta$ ,  $\phi$ , and  $\gamma$  are unknown parameters;  $\varepsilon$  is a stochastic term representing unmeasured influences; and  $\tau_j$  is a teacher fixed effect that provides a measure of teacher value added for teacher  $j$ .<sup>6</sup>

Statistical analysis is indispensable but can be very misleading unless supported by a valid theory of the underlying causes of the results. But, in fact, the results themselves cry out that something is amiss, since the value-added principle has exhibited far more

Not being curriculum based, reading tests cannot be accurate measures of school-driven gains in a given year.

Two false assumptions underlie applying VAMs to reading tests. The first mistake is the assumption that reading comprehension is a general skill. The second is the assumption that existing reading tests can accurately gauge the value that has been added by the teacher to reading comprehension from one year to the next. Our current reading tests cannot in fact reliably and validly gauge the value the teacher has added.

Here's why. Scores on reading tests reflect knowledge and vocabulary gained from all sources. Advantaged students are constantly building up academic knowledge from both inside and outside the school. Disadvantaged students gain their academic knowledge mainly inside school, so they are gaining less academic knowledge overall during the year, even when the teacher is conveying the curriculum effectively. This lack of gain outside the school reduces the chance of low-socioeconomic-status (SES) students showing a match between the knowledge they gained in school during the year and the knowledge required to understand the individual test passages.<sup>7</sup> The tests are fairly accurate means of gauging a student's general knowledge, but they have no way of indicating the sources of students' general knowledge. Not being curriculum based, they cannot be an accurate means of testing how well the particular knowledge

of 29 majors generally associated with four-year degree programs. The test scores of students seeking to enter graduate education programs are similarly low and, on average, undergraduate education majors score even lower than the graduate education applicant pool as a whole. Education schools long have accepted under-qualified students, then offered them programs heavy on pedagogy and child development and light on subject-matter content.

This scientific-sounding comment is incorrect from the start. The assertion that "Teacher quality affects student learning more than any other school-based variable" is not footnoted. According to two summaries of research by Russ Whitehurst, a better curriculum can range from being slightly to dramatically more effective than a better teacher.<sup>5</sup> That's not surprising when you consider that the curriculum is what teachers teach and what students are supposed to learn. Teachers are not to blame for ideas and curricula that are inherently inadequate.

Some policymakers have recently decided that the way to improve teacher effectiveness is to institute value-added teacher evaluations as part of a system of incentives, rewards, and sanctions, potentially including dismissal. The theory is that such a system will energize teachers, boost their performance, and bring highly qualified people into the profession. Some jurisdictions, including Chicago, Washington, D.C., and New York City,



in the school curriculum has been imparted. The implicit assumption that “general reading skill” is itself the content of the curriculum is a technical mistake and an incorrect assumption. Once that mistake has been exposed, the validity of the VAM projects in language arts collapses. Any judge in a lawsuit, properly alerted to the falsity of their assumptions, should rule against the fairness of value-added measures for rating language arts teachers. These reading tests may be roughly accurate measures of a student’s average reading abilities, but, not being curriculum based, they *cannot* be accurate measures of school-driven gains in a given year.

In short, there’s no valid or reliable way of determining what test-relevant verbal knowledge is school based and what is not. How could it be determined? *Tests that are curriculum-blind cannot gauge how well a curriculum has been imparted.* VAMs

Good teaching can often depend more reliably on the coherence of the wider system, and the cooperation it brings, than on virtuoso performances.

in reading are thus inherently unfair both to low-SES students and to their teachers. Reading tests at best are 70 percent accurate at the individual level.<sup>8</sup> The inherent uncertainty of the school-based contribution to a student’s reading scores between one year and another must reduce the validity of test inferences even more. Statistical manipulations cannot make a test reveal what it cannot reveal in principle. The whole VAM effort in reading will need to meet this objection head-on in order to establish the effort’s validity. It’s hard to see how it could do so. It has not done so thus far.

If I were a principal in a primary school, I’d spend my money on teachers, on their ongoing development, and on creating conditions in which the work of teachers in one grade supports the work of teachers in the next, and in which teachers would have time to consult and collaboratively plan. One especially vivid story about collaboration in Japanese elementary schools\* was told to me directly by the late professor Harold Stevenson, who studied Asian schools. He had observed the event in a fourth-grade math class. A student was having grave difficulty with a math problem and its concepts. After allowing the student to work on it for a short time, the teacher quietly made a surprising analogy with the student’s daily experience

as a way of dealing with the problem. The student’s face brightened, and he instantly began to solve the problem.

After the class, Stevenson went to the teacher to congratulate her (in perfect Japanese) on the most remarkable bit of teaching he’d ever witnessed. The teacher shook her head: no, it wasn’t her brilliance that produced the result, and from her desk drawer she took out a handbook that teachers had cooperatively compiled. “Here it is,” she said. “It’s suggested as a good tack to try when you run into that situation.”

The incident illustrated how good teaching can often depend more reliably on the coherence of the wider system, and the cooperation it brings, than on virtuoso performances. Schooling takes 12 years. Its success depends on slow but sure progress, not bursts of brilliance—welcome as those are when talented teachers inspire a whole class. □



## Endnotes

1. But there have been laudable triumphs in regular schools too. See E. D. Hirsch, Jr., *Why Knowledge Matters: Rescuing Our Children from Failed Educational Theories* (Cambridge, MA: Harvard Education Press, 2016), 159–184.
2. Philip Gleason, Melissa Clark, Christina Clark Tuttle, Emily Dwoyer, and Marsha Silverberg, *The Evaluation of Charter School Impacts: Final Report* (Washington, DC: U.S. Department of Education, 2010).
3. See, for example, Kate Walsh, Deborah Glaser, and Danielle Dunne Wilcox, *What Education Schools Aren't Teaching about Reading and What Elementary Teachers Aren't Learning* (Washington, DC: National Council on Teacher Quality, 2006).
4. A large part of human language interpretation is disambiguation, the process of choosing appropriate word and clause meanings, and rejecting others. Despite decades of work and billions spent, this problem of machine translation has not been solved. Yehoshua Bar-Hillel famously argued it could not be solved, in his piece “The Present Status of Automatic Translation of Languages,” *Advances in Computers* 1 (1960): 91–163. I have not seen a credible refutation of his argument, which is based on the insight that an unstated context is required for disambiguation. So far, no way has been devised even in principle to enable a machine reliably to identify which unstated context is the right one. Computers need explicitness; they seem to be very literal minded. So far, they are less expert than people in gauging the unsaid that is necessary to grasp the said. Moreover, they cannot come up with new meanings for old words—which humans do all the time. Landauer’s “Latent Semantic Analysis” makes a stab at analyzing what other words are and are not present, as does Google Translate (a good stab—but unreliable). See Thomas K. Landauer and Susan T. Dumais, “A Solution to Plato’s Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge,” *Psychological Review* 104 (1997): 211–240.
5. Matthew M. Chingos and Grover J. “Russ” Whitehurst, *Choosing Blindly: Instructional Materials, Teacher Effectiveness, and the Common Core* (Washington, DC: Brookings Institution, 2012); and Grover J. “Russ” Whitehurst, “Don’t Forget Curriculum,” *Brown Center Letters on Education*, October 2009.
6. Eric A. Hanushek and Steven G. Rivkin, “Generalizations about Using Value-Added Measures of Teacher Quality,” *American Economic Review* 100 (2010): 267–271.
7. The claim by test makers that their questions are self-contained or made fair by glosses is convenient but erroneous and naive. No text—glossed or not—is self-contained.
8. That is the average intercorrelation between the most reliable tests. Leila Morsy, Michael Kieffer, and Catherine Snow, *Measure for Measure: A Critical Consumers’ Guide to Reading Comprehension Assessments for Adolescents* (New York: Carnegie Corporation of New York, 2010).

\*For more about lesson study, see “Growing Together,” in the Fall 2009 issue of *American Educator*, available at [www.aft.org/ae/fall2009/dubin](http://www.aft.org/ae/fall2009/dubin).