Teacher Quality

Edited by Lance T. Izumi and Williamson M. Evers

(Hoover Press, 2002)

http://www-hoover.stanford.edu/publications/books/teacher.html#toc

TABLE OF CONTENTS

Teacher Quality

Eric A. Hanushek

Dr. Eric Hanushek, the nation's foremost education economist, addresses one of the most important questions in education policy—by what means can government improve the quality of the nation's teacher force, and how can it accomplish that goal without making the current problems worse? Hanushek's basic answer is that government should not prescribe solutions for local schools, but focus instead on providing incentives. According to Hanushek, "If the objective is to improve student performance, student performance should be the focal point of policy."

In the past, government has relied on regulations that determine education inputs, for example, class size and credentials. Rather than boosting student performance, these mandates have often had perverse effects. Shrinking class size increases the demand for teachers, while credentialing requirements, which do not ensure quality, limit the supply of candidates. The combined effect is that the teachers the school districts end up hiring are often low-performing.

Other variables, like differences in teacher ability, have far greater impact on student performance yet have largely been ignored by government. For example, teachers who elicit academic gains from their students are not rewarded for their achievements. Most teachers are hard-working and doing the best they can, but in the absence of incentives to improve, additional resources are not directed to maximizing student output. Hanushek argues that the adoption of performance incentives, while also holding schools and teachers accountable for the choices they make, is crucial if student achievement is to improve.

School reform is a topic on many people's minds today, and the air is full of advice and recommendations. Unlike many policy areas, the vast majority of people have strongly held opinions, mostly arising from their own personal experiences in school. As a result, much of policy making involves walking a line between research findings and popular views. Unfortunately, these popular views frequently are not the best guide for decision making.

This discussion begins with some evidence about the importance of teacher quality and moves to ideas about how the quality of teachers can be improved. Central to all of the discussion is the relationship between incentives and accountability. In simplest terms, if the objective is to improve student performance, student performance should be the focal point of policy.

From a policy perspective, although the proper role for different levels of government has been controversial, I believe that there are important things to be done by the federal government. These things are, nonetheless, quite different from both the current activities and many of the things that are being discussed.

THE IMPORTANCE OF TEACHER QUALITY

Starting with the Coleman Report, the monumental investigation in 1966 by the Office of Education, many have argued that schools do not matter and that only families and peers affect performance. Part of this view is true, and part is quite wrong. This report was the most extensive investigation of schools *ever* undertaken. Unfortunately, that report and subsequent interpretations of it have generally confused "measurability" with true effects. Specifically, characteristics of schools and classrooms, like the teacher having a master's degree or the class size being small, did not show any effect on student performance—leading to the conclusion that schools do not matter. This conclusion, probably more than anything else, led to a prevailing view that differences among schools are not very important.

The extensive research over the past 35 years has led to two clear conclusions. First, there are very important differences among teachers. This finding, of course, does not surprise many parents, who are well aware of quality differences of teachers. Second, these differences are not captured by common measures of teachers (qualifications, experience, and the like). This latter finding has important implications that I sketch below.

The magnitude of differences in teachers is impressive. Let me provide two different indications of teacher quality. For these measures I use a simple definition of teacher quality: good teachers are ones who get large gains in student achievement for their classes; bad teachers are just the opposite. Looking at the range of quality for teachers within a single large urban district, teachers near the top of the quality distribution can get an entire year's worth of additional learning out of their students compared to those near the bottom. That is, a good teacher will get a gain of one and a half grade-level equivalents, whereas a bad teacher will get a gain of only half a year for a single academic year. Alternatively, if we look at just the variations in performance resulting from differences in teacher quality within a typical school, then moving from an average teacher to one at the 85th percentile of teacher quality would imply that the better teacher's students would move up more than 7 percentile rankings in the year.

We can also return to the popular argument that family background is overwhelmingly important and that schools cannot be expected to make up for bad preparation from home. The latter estimates of teacher performance suggest that having three years of good teachers (85th percentile) in a row would overcome the average achievement deficit between low-income kids (those on free or reduced-price lunch) and others. In other words, high-quality teachers can make up for the typical deficits that we see in the preparation of kids from disadvantaged backgrounds.

Unfortunately, the current school system does not ensure any streaks of such high-quality teachers. In fact, it is currently as likely that the typical student gets a run of bad teachers—with the symmetric achievement losses—as a run of good teachers. Altering this situation is the school policy issue, in my mind.

CERTIFICATION AND OTHER CENTRAL APPROACHES TO QUALITY

In recognition of the importance of quality teachers, a variety of recommendations and policy initiatives have been introduced. Unfortunately, the currently most popular ones are likely to lower teacher quality rather than improve it.

The idea that has been picked up by policy makers at all levels is to increase the requirements to become a teacher. The idea is simple: if we can insist on better prepared and more able teachers, teacher quality will necessarily rise and student performance will respond. This argument—at least as implemented—proves as incorrect as it is simple.

The range of options being pushed forward include raising the course work requirement for teacher certification, testing teachers on either general or specific knowledge, requiring specific kinds of undergraduate degrees, and requiring master's degrees. Each of these has surface plausibility, but little evidence exists to suggest that these are strongly related to teacher quality and to student achievement.

More pernicious, these requirements almost certainly act to reduce the supply of teachers. In other words, the proposed requirements do little or nothing to ensure high-quality teachers, and at the same time, they cut down on the number of people who might enter teaching. Teacher certification requirements are generally promoted as ensuring that there is a floor on quality, but if they end up keeping out high-quality teachers who do not want to take the specific required courses, such requirements act more like a ceiling on quality.

The story on teacher certification initiatives is actually just a special case of a larger set of misguided policies that go under the name "input policies." These are attempts to specify pieces of the educational process. The recent craze for lowering class size—two years in a row the federal budget was held up until agreement could be reached on federal funding for hiring new teachers so that class sizes could be reduced—is the clearest example of an input policy: a variety of motivations have pushed this policy, which has little chance of success in terms of student achievement. This actually typifies the most common kinds of policies that we have been undertaking for the last three decades at least.

The Evidence on Inputs

The evidence on each of the input policy issues comes from a variety of sources but is very consistent. The simplest version is that we have been pursuing these policies for decades, and they have not worked. Table 1 shows the pattern of resources devoted to U.S. education since 1960. There have been dramatic increases in just the resources that people today advocate supplying. If we concentrate on the period of 1970 through 1995 (because we have student performance measures for a comparable period), we see that pupil-teacher ratios have fallen by

TABLE 1 Public School Resources in the United States, 1960–1995

Resource	*Pupil-teacher ratio*	*Percentage of teachers with master's or other higher degree*	*Median years of teacher experience*	*Current expenditure/ADA (1996–1997 dollars)*
1960	25.8	23.5	11	$2,122
1970	22.3	27.5	8	$3,645
1980	18.7	49.6	12	$4,589
1990	17.2	53.1	15	$6,239
1995	17.3	56.2	15	$6,434

close to a quarter, the number of teachers master's degrees has more than doubled, and median teacher experience has almost doubled. Because each of these inputs costs more, average real spending per pupil has increased by more than 75 percent, that is, by three-quarters after allowing for inflation. But if we look at student performance on the National Assessment of Educational Progress, we see that performance is virtually unchanged in math and reading and has fallen in science. This is hardly what the proponents of increased resources suggest should have happened.

This evidence on resources and performance is supported by detailed econometric studies. These statistical analyses of what goes on in the classroom provide little reason to believe that input policies will systematically improve student outcomes. While some studies suggest positive relationships with added resources, they are balanced by studies that actually show negative relationships. The existence of some positive findings allows advocates of specific policies to point to highly selective evidence supporting their cause, but it does not make for a different reality.

Similarly, with the recent push for class-size reduction, considerable attention has been focused on the Tennessee experiment of the 1980s, Project STAR. A much larger amount of uncertainty surrounds the evidence from this than most advocates want to acknowledge. Without going too far afield here, suffice it to say that Project STAR has been hugely overinterpreted. The clearest indication from this experiment is that very large reductions in class size (from 23 to 15) lead to small effects on student performance in kindergarten—hardly the evidence needed to support small reductions in class size at all grade levels.

The Policy Implications

It is important to understand how pursuing the conventional input policies could actually hurt the situation. As pointed out, increasing the requirements for teacher certification

could limit the supply of potential teachers and could thereby actually lower the quality of the typical teacher who ends up in the classroom. Similarly, lowering class size could hurt in two ways. First, it is very expensive, so it absorbs funds that could be applied to productive policies. Second, it expands the demand for teachers and can lower student achievement if the quality of new teachers ends up lower. Note, however, that we do not know much about the overall effects. The California class-size-reduction policy of 1997 indeed drew in more teachers who were not fully certified, but whether they were lower quality is unclear because certification is not closely related to effective performance in the classroom.

The generic issue is whether or not higher levels of government can effectively improve schools through uniform funding or with rules for how education is to be conducted in local schools. Here the evidence is quite clear. We do not know how to identify a well-defined set of inputs that is either necessary or sufficient for ensuring high-quality schooling. Finding such a set has been the Holy Grail of education research, and the search has been quite unsuccessful. Indeed, I do not believe that it is an issue of just needing more or better research. I simply do not think that we will identify (at least within our lifetimes) such a set with any clarity. I believe that the educational process is much too complicated for us to uncover a small set of criteria that are amenable to central legislation and control.

The evidence also underscores an aspect of the policy-making problem. Class-size reductions have been politically very popular. The federal government was merely mimicking the popular 1997 actions of the state of California. A large part of the political sentiment emanates from the commonsense arguments that persuade the general public that these are sensible policies. They just conflict with the evidence. And they imply that the policy maker must deal with political problems as well as policy problems.

PERFORMANCE INCENTIVES—AN ATTRACTIVE ALTERNATIVE

The simple position taken here is: *if one is concerned about student performance, one should gear policy to student performance*. Perhaps the largest problem with the current organization of schools is that nobody's job or career is closely related to student performance. Relatedly, popular input policies do nothing to change the structure of incentives. The key to effective policy is turning to performance incentives for teachers and other school personnel.

This is not to say that teachers or other school personnel are currently misbehaving. I personally think that most teachers are very hard working and that the vast majority are trying to do the best they can in the classroom. It is simply a statement that they are responding to the incentives that are placed in front of them (just as we all do). So when various decisions are being made, such as how to deal with added resources, the decisions may or may not be directed at the use that would maximize student learning. Instead, they might be directed at things that are publicly popular or things that make the decision makers' job easier or more pleasant.

The problem that goes along with this position statement is that we do not know the best way to structure incentives. We have not tried many performance incentive systems, so we have very little experience with or evidence from them.

A variety of approaches have been suggested and have conceptual appeal: merit pay for teachers, rewards to high-performing schools, and various forms of choice, including charter schools, tax rebates, and vouchers. Although evidence is slowly accumulating, the range of experiences is very limited.

There are nonetheless some things that we are quite certain about in the design of incentive structures.

Accountability and Value Added

One reason for general teacher resistance to incentive systems like merit pay is concern about what is being rewarded.

We know that families make a huge difference in the education of students. An implication of this is that we should not reward or punish teachers for the education they are not responsible for. If some students come to school better prepared than others, their teachers should not receive extra rewards. Similarly, if students come from disadvantaged backgrounds that leave them less well-prepared for school, we should not punish their teachers.

We want to reward teachers for what they add to a student's learning, that is, for their value added to the education of the child. Rewards should be geared to what teachers control, not to the specific group of students that they are given.

Pursuing this approach requires an aggressive system of performance measurement. We have to be able to track the progress of individual students, and we have to be able to relate this progress to the teachers who are responsible for it. This does not necessarily mean that we want a system of individual rewards as opposed to group rewards for teachers in a school, but it does mean that we have to accurately measure the performance of schools. This area—designing accountability systems—is an obvious area for federal leadership (although not necessarily for federal control).

Local Decision Making

It is also almost inconceivable that we could run a good performance incentive system from the national capital or even a state capital. If we try to devise the one best system and force it on local districts and schools, we will almost certainly fail. This statement really bites strongest when thinking about the limits of the federal government. Whereas the federal government can help provide funding for and guidance on the use of performance incentives, it is not in a good position to determine the "how" of the performance incentives.

At the same time, we should not simply assume that local districts and schools are currently able to make good

decisions. Personnel have not now been chosen for their ability to operate and manage different incentive systems. And, as mentioned, we do not have sufficient experience to provide any detailed guidance. Nonetheless, preparing local officials for performing these tasks is where we should be headed.

Neither should we assume that all policies that emphasize student outcomes and that provide performance incentives are altogether good. The design of incentives is complicated because many incentive structures lead to unintended and undesirable consequences. For example, if a move to broaden school choice led to complete racial or economic segregation in the schools, we would not think that it was a desirable policy. Therefore, we need to develop more experience with incentives and to evaluate these experiences. With incentive systems, the details generally prove to be critical.

LEARNING ABOUT INCENTIVES

In my opinion, one of the largest problems with education policy is that we never learn much from the policies we put into place. In fact, we frequently make policy decisions in ways that defy *ever* learning about their effects. The California class size initiative is a good example. All districts in the same state were simultaneously given financial incentives to reduce class size. Thus, even if one looks at student performance around the state, it is not possible to see what would happen in the absence of these incentives. Similarly, England recently introduced a broad policy of merit pay for teachers, but they did it everywhere at once. If student performance changes, is it because of the new incentives or because of other factors?

I realize that it is not the kind of policy that brings immediate political gratification, but I believe that nothing would have a more powerful influence on student performance ten years from now than a broad program of educational experimentation. The parallel with medicine is painfully obvious. In medicine, we are willing to admit that we do not know

everything about different procedures or therapies, and we conduct random-assignment controlled experiments to identify the effectiveness of different approaches. The results of this on the overall health of our population are clear and obvious. We also have a long history of social experimentation in health, welfare, and housing. We have learned an enormous amount over time that has helped to improve public policies. Nothing similar has occurred in education.

Experimentation and evaluation are legitimate federal roles. All states learn from these efforts, and no state takes into account the fact that evaluation results are useful to others. Without federal involvement there is likely to be too little investment in evaluation and knowledge production. Let me emphasize, however, that federal information collection is not the same as federal control of the schools, and there is no reason to expect that more centralized decision making would result from the federal government taking on a leadership role.

The problem, of course, is that experimentation and educational evaluation are not policies with mass appeal. Nonetheless, if we are to weed out bad policies and replace them with good policies, we need to accumulate evidence about performance.

CONCLUSIONS

Let me summarize.

1. Teacher quality is the key to improved schools.
2. Teacher quality cannot be readily linked to teacher characteristics; therefore, new and more extensive certification and training standards are unlikely to be effective.
3. Policies aimed at student performance instead of inputs offer the only real hope for improvement. Input policies, even though frequently popular, need to be resisted. At the same time, developing good accountability

systems is central, and the federal government can provide leadership (without nationalizing the process).

4. The federal government should limit its role to concerns of equity and of knowledge and should not attempt to act like a local school board. At the same time, the federal government should require performance for funds it disperses, such as the Title 1 funds that aid the education of disadvantaged students.
5. Developing improved policy requires better information about what works, and the most effective way of accumulating this evidence is the design of systematic experiments and evaluation.