Why Is Education Reform so Hard?

Eric A. Hanushek*
Stanford University and National Bureau of Economic Research

February 2002

* This paper was originally presented at Louisiana State University during a visit as a Reilly Distinguished Public Policy Fellow at the Reilly Center for Media and Public Affairs.

Why Is Education Reform so Hard?

By Eric A. Hanushek

Reform has been the operable word in education for many decades. We are continuously undergoing moves to reform, whether because the Russians launched the first satellite, because we wish to cure the problems of poverty, or because we feel are spending is not bringing desired results. Unfortunately, while reform of schools may well be appropriate, the approaches we have taken have been ineffective – and are likely to continue being ineffective unless significant changes are made.

This discussion begins by setting out the need for reform. It then outlines the alternative approaches to reform that exist. Finally, it provides recommendations for actions based on the ideas of increasing the performance incentives in schools and providing better information about the performance of schools.

The need for reform

Understanding the current state of U.S. education is important both for motivating reform discussions and for diagnosing possible reforms. The story begins with the pattern of student performance. Figure 1 tracks achievement of 17-year-olds over three decades. These data come from the National Assessment of Educational Progress, or NAEP, and are generally recognized as a reliable indicator of how performance compares over time. From this figure, reading and math performance are slightly up, while science and writing

are significantly down. In simplest terms, student performance looks flat over the last three decades of the twentieth century.

******** *Figure 1 here* *********

Being flat might not be bad if in fact this performance was high. But, by international comparisons, it is hard to say that this performance is particularly good. Table 1 displays the position of the U.S. on the Third International Mathematics and Science Study (TIMSS), an international test taken in 1995 by a large number of developed and developing countries. The performance of U.S. students, while near the center of the science distribution in the eighth grade, falls to the bottom grouping by twelfth grade. This performance actually came as no surprise, because the United States has performed similarly on each of the international tests that have been given since the early 1960s.

********Table 1 here *********

One observation at odds with these comparisons is the fact that the U.S. led the world in economic growth during the twentieth century, and most economists believe that education and human capital is an important part of growth. Part of this appears to be simply that the U.S. has substituted greater amounts of schooling for quality. Thus, by having higher levels of school attainment, we have managed to get around the lower content of each grade. Three other possibilities could also contribute to explaining our

better than expected economic performance. First, the U.S. higher education system, frequently regarded as the best in the world, is able to overcome the poorer training in primary and secondary schools. Second, even though the U.S. education system does not provide general skills to the population, it could encourage and develop creativity – creativity that comes through in inventions and technological change. Third, the exceptional growth performance of the U.S. economy could reflect factors other than schooling – an open labor market that encourages adjustment to new technologies, a well-oiled capital market, an economy relying more on private decisions than governmental decisions, and so forth. While we cannot currently distinguish among these explanations for the performance of the U.S. economy, I suspect that all contribute.

But, just because all enter does not mean that we can ignore the performance of our students and our schools. Other nations of the world, for example, are rapidly increasing the educational attainment of their populations. They also appear to be doing this in many cases without compromising the quality of their schools. Thus, the ability of the U.S. to substitute high quantity of schooling for lower quality in comparison to other countries is much less possible today than it was even twenty years ago.

There is another important aspect of the flat performance of U.S. schools: it does not reflect lack of trying. The U.S. has substantially increased the resources devoted to schools. Figure 2 shows the increase in real spending per pupil between 1890 and 1990. Over this period, spending per pupil grew at 3 ½ percent each year *after allowing for inflation*. This is a tremendous growth rate to sustain for a century.

******** *Figure 2 here* *********

Moreover, in the relevant recent period this change was accomplished in exactly the ways commonly advocated. Table 2 shows the decreases in pupil-teacher ratios and the increases in teacher degree levels and teacher experience from 1960 on. These elements each contribute importantly to the cost of schooling, so it is not surprising that the real spending calculations at the bottom of the table show a tripling between 1960 and 1995.

******** *Table 2 here* *********

These data suggest that the resources available to school have not been used in very productive ways. Increases in spending have not translated into increases in student performance.

Other factors can explain part but not all of this. Specifically, outside changes have put cost pressures on schools, implying that the education provided in the past may be more expensive to achieve today. First, kids may come to school less well prepared today than in the past. We know, for example, that the prevalence of single-parent families has increased over time. Tied to this, poverty rates among children have grown. These would be expected to work against student achievement – and would imply that schools have to work harder to achieve the same results now as in the past. But, offsetting these, the education level of parents has increased, and family sizes have fallen. These would be expected to work in favor of student achievement. Although it is

difficult to weigh the opposing factors with any precision, the best guesses suggest that they about evenly balance each other.

Second, other cost factors enter. Since the incorporation of requirements for serving handicapped students, the special education population has grown substantially: from eight percent of students in the late 1970s to 14 percent by the late 1990s. School programs for special education students on average cost more than twice those for regular education students. Similarly, there has been growth in students from non-English speaking backgrounds, entailing other programmatic expenditures.

Third, other industries have been competing for college graduates – particularly women – with increasing vigor. The salaries of college graduates have risen dramatically, meaning that schools must pay more for teachers just to have a chance at the same group of graduates that they did in the past.

Each of these latter factors does in fact contribute to the cost of school operations. Nonetheless, by any reasonable accounting, these factors explain only a portion of the increases in spending on schools. If these factors were removed from the calculations, substantial growth in real school spending would remain.

The aggregate trends are also confirmed by a large amount of detailed statistical study of what goes on in classrooms and schools. These studies, which attempt to uncover the effects of differing amounts of resources on student achievement, provide no evidence of a consistent or systematic impact of added resources. To be clear, some studies suggest that resources help, others actually find that more resources student achievement, but most provide no confidence that there is any relationship between resources and student performance.

These results do not say that money never matters. Nor do they say that money cannot matter. On the other hand, they indicate quite clearly that unless other more fundamental aspects of schools change increasing resources cannot be expected be expected to lead to noticeable improvements in student performance. Thus, these results reinforce the previous trends, where increases in overall resources did not correlate with changes in student outcomes.

In short, there is a clear need for reform. Desires to improve student outcomes in the past have not been realized, even with substantial injections of resources into the system.

Issues in improving student performance

Three factors appear to be extraordinarily important in explaining the current situation. First, there are no significant incentives within the system that push toward higher student achievement. Second, a variety of vested interests operate to distort decisions. But, third, and I think often neglected, people like their schools. I will quickly describe these ideas and then move to their implications for reform.

Incentives. Even though we are interested in raising the achievement of students, student performance has little or nothing to do with the pay, employment, or careers of most people in schools. To be sure, many people have chosen teaching careers because they want to help children, and they may take some real satisfaction from seeing students learn. Yet, almost everybody is motivated by a complex set of factors. In a situation where rewards are not aligned with the outcomes we are looking for, it should not be surprising that student performance is little moved by the series of reforms, programs,

and resources that we have directed toward schools. Another aspect of the lack of incentives relates directly to efficiency of schools. It is not possible to identify anything that looks like direct incentives to conserve on resources or, more importantly, to use resources efficiently.

Vested interests. Many of the people making decisions about school policy have a direct interest in the policies themselves, regardless of their impact on student performance. Thus, for example, teachers' unions have interests in increasing their membership – say through lowering class sizes – and these interests come independent of any potential impact on student outcomes. Similarly, the teachers themselves find that fewer students imply less work. Building administrators find that cooperative teachers who volunteer for a variety of activities outside of the classroom make their jobs easier. Such personal interests can in fact influence the kinds of policies and operational decisions that are made.

People like their schools. Regular surveys of attitudes toward schools reveal that everybody likes their own school. The median person thinks that his or her school gets a grade of B+. But, interestingly, those same people think that schools other than their own are, perhaps, graded at a C- or below. The disjuncture between the two is interesting and informative. The grades for schools by a random sample of people should provide a measure of the quality of schools in the country, i.e., of the schools other than each person's own school, but they do not. This is a classic case of Lake Wobegone, where everybody is above average. Yet this misperception by parents is not innocuous, because families are induced to protect the status quo, their good school.

These three factors have had an influence on the character of reforms that we have seen in the past, and they are likely to influence what happens in the future.

Past reforms

The general character of past reforms is best described as regulatory in nature, supplemented by ever increasing resources. In recognizing the need to improve, it has been natural for policy makers in the legislature, in state education departments, and in local school districts to attempt to make schools work better. A major portion of this has been guarding against really bad things. For example, putting regulations on what is required for teacher certification can be viewed simply as declaring the minimum quality of teachers that can enter the classroom. Alternatively, prescribing the process for diagnosing educational handicaps and developing an educational plan can be seen simply as ensuring that needy students are served at an acceptable level by local school districts.

This approach does provide a series of coarse incentives, because there are explicit or implicit sanctions related to violating these regulations. But, whether or not they are incentives that improve student performance depends crucially on whether the regulations incorporate factors that are systematically related to student achievement. Unfortunately, the available research suggests that most of these regulations induce outcomes that are *not* related to student achievement.

This regulatory environment in fact may in many instances have the opposite effect of that intended. Take, for example, the introduction of tougher certification requirements. If these requirements do not lead to better teachers in the classroom, then they cannot have beneficial effects. They can nonetheless have harmful effects if, on the

other hand, they lead potentially good teachers not to enter teaching. Such could be a natural result whenever the regulations make it more difficult and time consuming – that is, more expensive – to enter teaching.

Given our current knowledge, it is very difficult to design regulations that effectively improve the quality of schools. Doing such requires detailed knowledge of the determinants of student performance, and we do not have the requisite knowledge in a wide ranges of areas.

Current reform

The recent period has seen a change in the direction of much of the reform of schools. Instead of focusing on the various inputs to schools, attention has switched more to goals and outcomes of schools. Specifically, "standards based reform" is designed to begin with a delineation of what we hope schools to accomplish. While this can be quite broad, it can also be very pointed and specific, such as what mathematical concepts we believe that fourth graders should master. It also includes the notion that we will measure movement toward meeting these goals, through testing individual knowledge and comparing the results of that to the goals.

Standards based reform is closely linked to ideas of testing and accountability. The recent expansion of interest in measuring student performance demonstrates the attention to these ideas. The annual student testing program in Texas that has been going on since 1993 is an example of linking standards, measurement, and accountability. The move of the New York State Regents to require all students to meet the demanding standards of the Regents Diploma is another. And, the recently enacted re-authorization

of the federal Elementary and Secondary Education Act calls for broad expansion of such ideas across all of the states.

Most such reform does not, however, provide a broad set of incentives that are linked to the accountability system. A number of states have discussed and/or implemented sanctions for students who do not meet the requirements. These high stakes components currently are more discussed or projected than binding. But, even if states follow through, there are confused incentives, and the incentives for teachers and schools are often muted or nonexistent.

Again, part of the problem is reflected in the complexity of the educational process. Schools are just one input to a student's learning. Families and friends are known to be a very important input. Similarly, the student may have limited control over the school experiences he or she faces, so the achievement they reach is partly affected by each student's effort but surely not entirely. From the school and teacher viewpoint, they must accept the preparation that a student has when they come to the school and to the specific classroom. Surely it would be inappropriate to hold the teacher responsible for the starting level of achievement of the students in a class.

The desirable aspects of accountability systems are real, but many current accountability systems send the wrong signals. These issues are likely to have significant impacts on the results that are seen.

The current reform efforts actually include another set of policies and activities. Most people now understand the importance of the teacher quality. In recognition of this, a variety of efforts have been undertaken to ensure high quality teachers in the classroom. The most popular approach involves tightening up on who teaches, that is, trying to make

sure that only the best trained and prepared potential teachers can be teachers. This approach generally involves a combination of mandating additional training requirements such as subject matter specialization along with a master's degree in teaching, higher entry scores on teacher examination, and obtaining a degree from a school of education that meets strict accreditation standards.

These reform efforts are in reality a continuation of prior regulatory approaches, and their value is directly linked to whether or not they induce better quality teachers in the classroom. Again, however, the things that enter into the certification requirements – no matter how intuitively appealing they are – rest on very little evidence that they are important. For example, past work suggests that having a master's degree or not is unrelated to a teacher's performance in the classroom. Thus, adding a requirement for a master's degree makes obtaining a teaching certificate more costly (thus reducing the group of people willing to prepare for teaching) with no commensurate improvement in skills that are likely to be translated into higher student achievement.

An alternative direction

The alternative approach returns to the general outline of aspects of schools that could explain our current state. At the outset, there seems to be no substitute for providing betters incentives to those in schools. *If we are interested in student performance, we should focus our attention on student performance.* Thus, people have to realize that their performance will be judged on the basis of student achievement and that judgments about their performance will translate into tangible results that affect them.

The tough part of course is that we have very little experience designing performance incentives for schools. Many different ideas have been introduced, but

schools have resisted any efforts to introduce and evaluate them in a systematic way. The potential incentives sometimes keep the basic structure of the existing public schools but add other elements. For example, the introduction of merit pay for teachers where rewards are directly linked to student performance operates within the current schools but could noticeably change incentives. Similarly, many states have enacted or proposed systems of rewarding entire schools based on the performance of students in those schools. Examples are found in California and Texas, where information on student achievement triggers rewards to schools.

Some appealing options, however, would yield more fundamental changes in the organization and structure of public schools. The clearest example is the introduction of more parental choice into the schools. While the least different form would allow students to choose among just existing public schools, more radical ideas like vouchers would represent substantial structural alterations. The idea behind these is that parental choice and student movement across schools triggers movement of resources, creating incentives for existing schools that do not want to lose resources.

These systems are each likely to work through their impact on the quality of teachers. If the systems create incentives for better student performance, they also create incentives to ensure that the schools hire and retain high quality teachers. As such, they supplant the efforts to tighten up on teacher entry and certification. If school personnel are to be held accountable for student results, they need to have the flexibility to make decisions that they believe will enhance outcomes. Indeed, the philosophy behind these reforms is the opposite of that behind the currently popular proposals. Instead of

tightening up on requirements it would be better to loosen up on requirements – but then to hold school personnel accountable for the student outcomes.

As mentioned, however, we have little experience with these alternative incentive approaches. Moreover, we know that the design of incentive systems is indeed complicated and difficult to do well without prior experience. It is likely that any newly developed systems will begin with undesirable components. Setting up desirable incentives without also introducing unanticipated and undesirable incentives is an art that often builds on prior mistakes. Thus, there are neither certain answers to the questions "what is the best set of incentives?" nor can we count on early attempts working smoothly. It would be a huge mistake, however, to use our current ignorance to maintain the status quo.

The history of the past, which has both avoided incentives and has failed to achieve success, underscores the fact that we cannot avoid trying different approaches. It also underscores that we must build in the idea of learning from any reforms that we try.

Pressures for the status quo

The current system contains significant pressures to maintain the status quo. First, the current personnel in the system, people who have significant influence on the course of educational policy and practice, generally do not favor any structural change in the system. They may prefer some new programs, some alternative ways of organizing the existing schools, and the like, but they seldom are strong advocates for increased accountability, for direct performance incentives, or for outside competition. These positions are easy to understand, even if they offer significant resistance to major reform.

But parents also resist change. After all, they like their current schools. Change in the organization and incentives of schools can only have the result of potentially disrupting their well-functioning school. Thus, for example, when Michigan voters in 2001 faced a referendum to permit vouchers for low-income families, middle-income voters resoundingly defeated it. One interpretation is simply that offering more options to low-income families held out the possibility of disrupting the schools of this "unaffected" group – so a risk averse response is just to stick with the status quo.

Uncertainty about change operates in the same direction. If you like your current school, an uncertain reform proposal does not look particularly appealing, because it has the potential of worsening the current situation.

Ideas on reform

Reform is needed. From a variety of perspectives, the current system is not performing up to hopes or expectations. Performance remains flat at levels that are uncompetitive internationally. And the system keeps absorbing greater resources with no commensurate improvement in student achievement.

The best hope for reform lies in improving incentives related to student outcomes. Pursuing this goal is aided by current movements to measure and to focus on student achievement, but accomplishment of the goal needs more than the current accountability. In some manner, it is necessary to relate resources to performance. A variety of proposals exist, but uncovering the best one for varying circumstances will require experimentation and evaluation.

A key element in obtaining reform, in my opinion, rests on providing very accurate information about performance. In particular, by the data and logic developed above, some parents are currently misled about the quality of their schools. The typical school of the nation is not B+ or better, but instead needs improvement. Yet parents today are not provided sufficient information about the value-added of their school to be able to assess their school quality accurately. First, student performance for many schools has not been generally available. Routine and regular testing and reporting of results on student performance have been restricted to a few states. This is changing, in part because of federal legislation. Second, however, it is often difficult to sort out the value-added of schools from the preparation of students. We know that families are very important in education, so it should come as no surprise that students in middle-class dominated schools perform well on many achievement tests. Yet, this does not necessarily signal a high value-added or contribution of schools to their achievement. An accountability system that sorted out these components and that reported on the quality of schools would come to a shock to many parents, because all schools would no longer be above average.

There is reason for hope. A few states have begun to publish student achievement information regularly for all students and grades. This information is sometimes, but not always, accompanied by other information that allows crude adjustment for the preparation of students (the input of families, abilities, and other things). We see in some instances (Florida, Texas, and North Carolina) that such information contributes to some dramatic performance improvements by schools at the bottom. The exact mechanism is unclear, but these schools respond to being put in the spotlight.

We have to pursue such innovation. The future well-being of our citizens depends importantly on this.

Table 1. Country Ranking of Performance on Third International Mathematics and Science Study (TIMSS), 1995

8th Grade Performance		12th Grade Performance	
Mathematics	Science	Mathematics	Science
Singapore	**Singapore**	**Netherlands**	**Sweden**
Korea	**Czech Republic**	**Sweden**	**Netherlands**
Japan	**Japan**	**Denmark**	**Iceland**
Hong Kong	**Korea**	**Switzerland**	**Norway**
Belgium-Flemish	**Bulgaria**	**Iceland**	**Canada**
Czech Republic	**Netherlands**	**Norway**	**New Zealand**
Slovak Republic	**Slovenia**	**France**	**Australia**
Switzerland	**Austria**	**New Zealand**	**Switzerland**
Netherlands	**Hungary**	**Australia**	**Austria**
Slovenia	England	**Canada**	**Slovenia**
Bulgaria	Belgium-Flemish	**Austria**	**Denmark**
Austria	Australia	**Slovenia**	Germany
France	Slovak Republic	**Germany**	France
Hungary	Russian Federation	**Hungary**	Czech Republic
Russian Federation	Ireland	Italy	Russian Federation
Australia	Sweden	Russian Federation	UNITED STATES
Ireland	UNITED STATES	Lithuania	Italy
Canada	Germany	Czech Republic	Hungary
Belgium-French	Canada	UNITED STATES	Lithuania
Sweden	Norway	*Cyprus*	*Cyprus*
Thailand	New Zealand	*South Africa*	*South Africa*
Israel	Thailand		
Germany	Israel		
New Zealand	Hong Kong		
England	Switzerland		
Norway	Scotland		
Denmark	*Spain*		
UNITED STATES	*France*		
Scotland	*Greece*		
Latvia	*Iceland*		
Spain	*Romania*		
Iceland	*Latvia*		
Greece	*Portugal*		
Romania	*Denmark*		
Lithuania	*Lithuania*		
Cyprus	*Belgium-French*		
Portugal	*Iran*		
Iran	*Cyprus*		
Kuwait	*Kuwait*		
Columbia	*Columbia*		
South Africa	*South Africa*		

Note: **bold** – significantly about United States; *italics* – significantly below United States.
Source: U.S. Department of Education (1999)

Table 2. Public School Resources in the United States, 1960-1995

Resource	1960	1970	1980	1990	1995
Pupil-teacher ratio	25.8	22.3	18.7	17.2	17.3
% teachers with master's degree or more	23.5	27.5	49.6	53.1	56.2
median years teacher experience	11	8	12	15	15
current expenditure/ADA (1996-97 $'s)	$2,122	$3,645	$4,589	$6,239	$6,434

Source: U.S. Department of Education (1997)