In his stimulating article, Eric Hanushek argues that school quality bears little relationship to expenditure on school inputs; that achieving quality should be given greater priority relative to quantity; and that schools should be given greater incentives to improve quality. In this comment, I argue that the limited available evidence suggests that expenditure affects school quality and that we can be fairly confident that supplying particular inputs, such as radio education and textbooks, will improve school quality. I also argue that we have insufficient evidence to conclude that quality should be a higher priority than ensuring that schools are available for more children. Finally, I add some cautionary notes to Hanushek’s call for increased emphasis on incentives for schools and teachers.

Despite these areas of disagreement, Hanushek and I may differ as much in temperament as anything else. We agree that investments in text books and radio education are worthwhile and that reducing class size is a lower priority. Whereas Hanushek (p. 234) interprets the fact that different studies have reached different conclusions as “fairly conclusive evidence that measured resources are not systematically related to student performance,” I read the weight of the evidence as indicating that increasing resources typically improves student performance modestly. We both support trying new incentive systems for schools.

The Effectiveness of Educational Expenditure

Hanushek surveys studies that examine the correlation between educational inputs and test scores and concludes that there is little evidence of a systematic relationship between the two. I argue that these studies are seriously flawed; that other methods of summarizing these studies suggest that increased spending for education does improve test scores; that some studies indicate that additional expenditures on education will increase students’ future wages; that randomized trials (an evaluation methodology), which are less subject to statistical problems, indicate high payoffs to particular education expenditures; and that
high levels of spending on education in the United States suggest that the market values educational inputs.

The studies Hanushek reviews examine the statistical correlation between school inputs and test scores. These studies interpret a positive correlation between an input and test scores—that is, a tendency for the schools with high test scores to have high levels of the input—as evidence that the input is effective. These correlations, however, may be affected by many factors other than the effectiveness of inputs. For example, if schools with better facilities tend to be located in communities that have a stronger commitment to education, the correlation between facilities and test scores may be positive, even if the facilities do not themselves increase test scores. Conversely, if it costs more to provide education in remote, underdeveloped areas than in cities and towns, there may be a spurious negative correlation between expenditures and test scores. As Hanushek notes, good teachers may attract and retain more students, inducing a positive correlation between class size and student performance. Hiring more teachers and reducing class size, however, could still increase test scores. Many of the studies attempt to control for factors that could lead to spurious correlations, but it is impossible to control fully for such factors.

Although Hanushek summarizes these studies as indicating little relationship between school inputs and test scores, different conclusions have been drawn from the same data. Hanushek counts the studies with positive, negative, and statistically insignificant results and argues that, if a large number of studies were statistically insignificant relative to the number that were positive, there is no strong evidence that the input is valuable. Due to measurement error and random difference in test scores, any study will have some probability of detecting a positive correlation even if there is no true effect, and some probability of failing to detect a positive correlation even if there is a true effect. But as Hedges, Laine, and Greenwald (1994) point out, the studies were designed to give a small probability of falsely detecting a significant effect of educational inputs but not to avoid the possibility of failing to detect such an effect if in fact it exists. Therefore, the chance of obtaining an insignificant result might be very large, even if the true correlation is positive. The chance of obtaining several significantly positive results is tiny, however, if the true coefficient is less than or equal to zero. For example, of the twelve studies of the effect of education expenditures per pupil, six found statistically significant positive results, six found statistically insignificant results, and none found negative results. Assuming that each study had only a 5 percent chance of yielding a significantly positive result if the true coefficient were zero, the chances of six or more studies out of twelve being significantly positive if the true effect were zero is only eleven in one million.

Table 1 reproduces Hanushek's table 1, with additional columns showing what the chances are of obtaining as many positive results as were found if the true effect were zero or negative (and the probability of obtaining as many negative results if the true effect were positive). If one accepts the studies at face value and assumes at most a 5 percent chance that a study would yield a signifi-
cantly positive coefficient if the true effect is zero or negative, there would be overwhelming evidence that all the inputs Hanushek examines, aside from the teacher-pupil ratio, have positive effects.

One difficulty with Hanushek's methodology is that it weights all the studies equally, despite differences in number of observations, procedures, and controls. Hedges, Laine, and Greenwald (1994) examine the same studies using a modern meta-analysis approach that corrects for some of these differences and uses more sophisticated procedures for aggregating the information in the different studies. They find a positive relation between spending on education and output. (For a review of the debate on meta-analysis, see Hedges, Laine, and Greenwald 1994 and Hanushek 1994.)

While Hanushek surveys the literature on the effect of education spending on test scores, other studies examine the effect of education spending on earnings later in life. Whether spending on education is a good investment on economic grounds depends on its effect on future earnings—not its effect on test scores. Card and Krueger (1992) find that migrants coming from areas in the United States with higher expenditures on schooling earn more in their destination city than migrants coming from areas with lower levels of spending. Card and Krueger (1994) survey a dozen studies on the effect of school expenditure on income. These studies, conducted over the past twenty years in industrial countries, typically show that a 10 percent increase in education spending is associated with a 1 or 2 percent increase in annual earnings later in life and that the net present value of increasing educational expenditures in the United States is likely to be positive. Although Card and Krueger's methodology has the advantage of examining the direct impact of education on earnings, it is important to note that

Table 1. Effects of School Inputs on School Quality

<table>
<thead>
<tr>
<th>Input</th>
<th>Number of studies</th>
<th>Number of studies*</th>
<th>Probability if true coefficient is negative</th>
<th>Probability if true coefficient is positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher-pupil ratio</td>
<td>30</td>
<td>8</td>
<td>8.46 x 10^{-5}</td>
<td>8.46 x 10^{-3}</td>
</tr>
<tr>
<td>Teacher's education</td>
<td>63</td>
<td>35</td>
<td>4.30 x 10^{-29}</td>
<td>0.83</td>
</tr>
<tr>
<td>Teacher's experience</td>
<td>46</td>
<td>16</td>
<td>3.58 x 10^{-10}</td>
<td>0.67</td>
</tr>
<tr>
<td>Teacher's salary</td>
<td>13</td>
<td>4</td>
<td>0.0031</td>
<td>0.135</td>
</tr>
<tr>
<td>Expenditure per pupil</td>
<td>12</td>
<td>6</td>
<td>0.000011</td>
<td>1.00</td>
</tr>
<tr>
<td>Facilities</td>
<td>34</td>
<td>22</td>
<td>7.26 x 10^{-21}</td>
<td>0.24</td>
</tr>
</tbody>
</table>

a. The number of studies in which the input has a statistically significant effect (positive or negative) on school quality.

b. The probability that this number of studies would record statistically significant positive effects if the true coefficient were actually negative or zero (third column) or statistically significant negative effects if the true coefficient were actually positive or zero (fifth column).

Source: Author's calculations.
the methodology is also subject to econometric problems and that others have found different results (see Heckman, Layne-Farrar, and Todd 1994; Betts 1994; and Hanushek, Rivkin, and Taylor 1995).

A third type of evidence suggesting that particular educational inputs can improve student performance is provided by the few randomized trials of educational interventions. Unlike retrospective studies, such trials are not subject to bias from omission of variables correlated with educational outcomes. The few existing randomized evaluations of educational inputs indicate that radio education and textbooks have substantial effects on educational outcomes and smaller classes have less effect. For example, Jamison and others (1981) randomly divided all primary schools in three provinces of Nicaragua into three groups: one group received radio mathematics education, another group received textbooks, and a third group served as a control. Radio instruction on mathematics was found to raise test scores by 1.2 standard deviations; textbooks raised the scores by a third of a standard deviation. In another study, Finn and Achilles (1990) report that the State of Tennessee randomly assigned students to classes of different sizes. Although some students transferred classes, which could bias the results, students in smaller classes scored one-twentieth of a standard deviation higher on tests.

A final source of evidence comes from the United States, where local school districts decide how much to spend on education. Mobility is high in the United States, and many people choose their communities, at least in part, on the basis of the quality of the school system and the level of local taxes. If additional spending produced no positive impact, people would move out of school districts that spend more money on education. But in fact, these areas record high property values, indicating that a significant proportion of the population values this level of expenditure. Spending on education is relatively high in the United States compared with that in other countries, suggesting that people in a decentralized market are willing to pay for education.

Hanushek argues that providing additional resources does not systematically improve school performance, and that, for policy purposes, cases where resources do improve performance are irrelevant, because such cases cannot be identified. I agree that the impact of additional resources varies with the circumstances, but I draw different conclusions. When I have a headache, I take aspirin because I know that on average it helps, even though aspirin only cures some headaches. Similarly, if I had to allocate resources centrally, I would try to provide textbooks to all schools, even though the textbooks might go to waste in some schools. Because different inputs are effective in different circumstances, it may be useful to decentralize spending on education. Still, as long as school systems are centralized, policymakers should allocate resources on the basis of the average impact.

Despite the preponderance of evidence indicating that increased expenditure on education typically improves students' future prospects, there is not enough evidence to judge whether the benefits of increasing overall education expendi-
ture outweigh the costs. The fact that many studies could not detect the effect of school inputs on test scores suggests that some inputs may have only a small impact.

A more important question is whether there are any areas in which additional investment in education will produce large payoffs—or where budgets could be cut at little cost. Here, Hanushek and I agree. Jamison’s experiments indicate that investments in radio education have tremendous payoffs and that, where textbooks are scarce, providing them to students will yield high returns. Hanushek and I also agree that funds should not be spent on reducing class size until higher-priority investments are undertaken.

I concur that little is known about what improves school quality and that additional nonrandomized studies are not likely to add to our information. Randomized trials do hold great promise, however (Newman, Rawlings, and Gertler 1994). In some cases, schools can be randomly divided into two groups, one of which receives a particular intervention. If this approach is not politically feasible, the two groups could receive the same intervention at different dates. Alternatively, two groups could receive different interventions. For example, additional textbooks could be issued to one set of schools, and the pupil-teacher ratio reduced at another. Schools in developing countries are especially well suited to randomized trials because significant changes in school budgets can be accomplished at very low cost. Nongovernmental organizations (NGOS) often assist individual schools, because they cannot afford to assist all schools, and randomization may be the fairest way of choosing which schools to assist. The NGOS may not be subject to the same political constraints that make it difficult for governments to undertake such studies. Randomized trials revolutionized medicine early in this century, and there is no reason they should not revolutionize education early in the next century.

Quantity versus Quality

Hanushek contends that policymakers should focus on improving school quality rather than quantity. On the basis of the wide range in annual gains in test scores among schools, he argues that school quality varies greatly. But this variation might just as easily reflect measurement error or the unmeasured characteristics of the surrounding community. The differences might also reflect the fact that some headmasters and teachers are better than others. This would be true even in an ideal system of school finance and does not imply that better policies could raise all schools to the level of the best schools.

Hanushek argues that improvements in school quality can pay for themselves by improving attendance and reducing repetition rates, thus reducing the number of years of schooling required to produce a student who has completed a given grade. He calculates that $1 invested in textbooks will yield a $10 return. I agree that investments in textbooks are a good idea, but his calculations im-
licitly measure inputs by the number of years a pupil attends school—even part
time and output by the number of grades the student completes. Because chil-
dren in developing countries have many responsibilities other than school, they
may, for example, attend school half time and complete fourth grade after eight
years in school. If the quality of schooling is improved, they may attend school
full time and complete fourth grade in only four years. These improvements
may well be warranted, but theoretically, the education system could still be
using approximately the same amount of resources in children's time, teachers'
time, classroom space, and supplies to produce fourth-graders. (In practice, fewer
resources will probably be needed because students will avoid repeating the same
part of the curriculum twice.) The efficiency of the system will not have doubled
any more than it would have if all the people taking college courses in the United
States at night and on weekends quit their day jobs and went to college full time
in order to graduate in half the time.

I agree that improving school quality is important and that it will increase
attendance. There is no evidence, however, that it is more important than opening
new schools in underserved areas or than subsidizing the cost of schooling to
allow more people to attend. As Hanushek notes, many studies have found high
rates of return to additional years of education, and, although many of these
studies are seriously flawed because of potential statistical problems, others are
quite convincing. Strong evidence indicates that spending money on textbooks
and on radio education will significantly improve school quality, but little evi-
dence is available about other inputs. Thus I think that, once textbooks have
been supplied at a ratio of one textbook for every two students, additional re-
sources would probably be better spent in ways that allow more children to
attend school. In some countries, that might mean opening new schools, and in
others it might mean subsidizing schools to make them more affordable.

School Reform

Although I agree with Hanushek that policymakers should experiment with
new incentive systems for schools and teachers, I would add some cautionary
notes. As he points out, there is, as yet, little evidence that these incentives will
improve education. Moreover, it may be difficult to design incentive systems
that do not lead schools to switch from teaching unmeasured objectives to teach-
ing objectives that are measured, to the detriment of students. For example,
Kenya's primary schools are judged by the results achieved on a national eighth-
grade exam. Many schools have responded to this incentive by promoting only
their best students to the eighth grade and forcing the others to repeat the sev-
enth grade. The schools are responding to incentives, but the incentives are too
narrow.

One way to provide broader incentives is to let parents choose schools through
a voucher system, based on whatever criteria they think are important. In indus-
trial countries, one of the main objections to voucher systems has been the fear that they would lead to segregation by race, class, and educational ability. The limited mobility and relatively low density in some rural areas of developing countries suggest that this may not be a problem; there may be room for only one school in a given area. However, the same low density may make education a natural monopoly, thereby providing a different rationale for it to be undertaken by the state because private providers might exploit their monopoly position.

Perhaps the main reason education is so widely under state control is that schools transmit values to children. Although attempts to instill values through the education system are sometimes seen as insidious, they may be to everyone's advantage. To take a trivial example, everyone would prefer that other people's children be taught not to litter, but they may not choose schools for their children on the basis of the strength of the anti-littering education. More seriously, one of the main challenges facing many developing countries is increasing national cohesion. If parents choose schools for their children, they might well choose schools that emphasize their particular regional, ethnic, or religious identity at the expense of other groups or of national identity. Although, theoretically, the government could subsidize the teaching of values by private schools, in practice, it might be difficult to enforce. The state can mandate that teachers repeat the pledge of allegiance and tell their students not to litter, but unless the state actually runs the schools, it is hard to stop the teachers from winking as they do so.

To conclude, there is considerable debate about the effects of increasing spending for education, although there is strong evidence that certain types of expenditures do improve student achievement. New incentive structures and organizational forms may provide better, less expensive education, but pitfalls should be expected along the way. Randomized evaluations provide the best way of obtaining more information about the value of different inputs and the effects of different incentive structures.

Notes

Michael R. Kremer is assistant professor of economics at the Massachusetts Institute of Technology.

1. Probabilities were calculated using the binomial formula.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.


