

THE QUESTIONS NO ONE HAS BEEN AFRAID TO ASK

Iris C. Rotberg

Thank you for inviting me this afternoon. I must say that I feel a little like a speaker on a panel discussing the Johnstown, Pennsylvania, flood with Noah sitting in the audience. All of you have had a lot of experience in this area, and many of you have asked me rhetorically and explicitly many questions about the subject matter of this forum. Therefore, I thought that it might be useful to title my remarks "all the questions no one has been afraid to ask." But I narrowed them down to 10. So, with apologies to David Letterman, let me get to the list--in no particular order of importance.

1. If the international test score comparisons all produce similar findings, doesn't that mean that there must be some underlying validity to them?

Not at all. It means that they all have the same shortcomings. First, for many reasons, the students represented in the test comparisons are much more highly selected in some countries than in others. Second, some countries, like the United States, have a relatively high proportion of low-income students who are in school and taking the test. I speak here not of absolute income but of the gap within various countries between the affluent and the poor. Countries with high numbers of low-income children who are still in school will be at a disadvantage on the test for reasons having little to do with the quality of the education system. Third, there are differences in curriculum emphases among nations that contribute to the relative rankings. If students haven't studied calculus, they will not do well on a calculus test. We don't need to administer an international test to tell us that. Moreover, the test scores won't help us decide whether or not more students should take calculus. That is a complicated question, and testing students who have never studied the subject won't help us answer it.

2. What made you question the findings of the international comparisons in the first place?

First, familiarity with some of the countries and education systems represented in the tests. Because of practical constraints on sampling, the studies did not reflect the realities in many of the

Iris C. Rotberg is a Senior Social Scientist with the RAND Corporation's Institute on Education and Training. This paper is based on informal remarks at a forum on the value of international assessments, presented at the Office of Educational Research and Improvement, United States Department of Education, March 23, 1992.

participating countries. Second, test score results that could not be explained by differences in the quality of education, for example, major reversals of rankings between higher and lower grades in Hungary, British Columbia, England/Wales, Japan, Hong Kong, the former Soviet Union, Slovenia, and the United States. Third, my experience with test patterns in general. In the United States, states' average SAT scores are highly correlated with the proportion of students who do not take the test. In commenting on this point, I suggested in an article--I thought facetiously--that ". . . one way to increase a state's average SAT score would be to discourage students from applying to colleges that require the test!" I received a letter asking whether I had used an explanation point because I believed that sort of thing couldn't happen. It turned out that the letter writer resided in an affluent and therefore high SAT district that had recently merged with a low-income and low SAT district. The unhappy development was that the students attending the schools in the low-income area were discouraged from taking the SAT in an attempt to keep the district's reputation--and property values--intact.

3. With all our expertise in statistics and sampling design, can't we simply improve the validity of the international comparisons?

No, we can't. The fact is that we have had expert statisticians working on the problems for the past 30 years. The difficulty is not in devising elegant statistical designs, but in carrying them out in the real world. The problems in making these comparisons are endemic to all of the studies, including the most recent ETS study which went out of its way to point out these problems and strongly advised in its press release against ranking the countries. But more important, would our children's education improve if we established rigid international controls on each nation's sampling design, located out-of-school (or homeless) children and tested them on science and mathematics, or controlled for tracking or relative socioeconomic status? And even if we did so, what is the chance that the test score differences could be attributed to the quality of each nation's education system?

4. Does it matter if we exaggerate the problem in the United States when we all agree that science education can be better than it is?

Yes, it does matter. First, the rhetoric is not supported by the facts. We incorrectly assume that adverse test score differences mean that our schools, or our parents, or our students, or our scientists, or our research institutions have failed. Second, policymakers, and students, make decisions based on this rhetoric. For example, a recent New York Times article describes the anger of young Ph.D. physicists who can't find jobs (evidently Amherst College received 813 applications for one position) amid government reports of "shortages." I am particularly concerned about proposed remedies based on misleading test score differences--for example,

raising course and graduation requirements--without doing anything about the vast financial differences between rich and poor school districts. Those requirements will do more harm than good. We are likely to screen out of the education system precisely those students who already receive the lowest quality education. We will reduce their graduation rates, and subsequent employability and earnings. We will end up with a so-called meaningful high school certificate, but fewer students will receive it. What will the others do for a living? Moreover, for all students, we are in danger of placing increased emphasis on rote learning, measured by multiple choice tests, and less emphasis on the type of curriculum changes and teaching practices that would focus on an understanding of basic scientific concepts and research methods.

5. Do you believe that we can learn something from other nations' education systems or teaching practices?

Of course we can. However, the challenge is to identify those practices that can realistically be replicated in the United States. While comparative international studies may provide some insights, relatively few produce findings that can be readily transferred from one nation to another. In most cases, it would involve a basic restructuring of a nation's social, cultural, and political structure. It would involve changes in some rather fundamental aspects of our society ranging from the respective roles of national and local governments in education, the role of the teacher in society, teachers' salaries, competitive sports in schools, summer vacations, our value system with respect to pluralism, open access to higher education across socioeconomic groups, the role of industry in vocational education and apprenticeship programs, and similar issues that each country looks at differently. The fact is we cannot--nor would it be wise to--superimpose changes in education outside the context of a country's cultural and political environment. That is not a problem limited to education. For example, comparisons of industrial policy in the United States and Japan are just as complex. Government/industrial links differ so fundamentally between the two countries that, as a practical matter, it remains unclear how the lessons learned in Japan might be transferred to the United States.

6. Aren't you being complacent about the problems in American science and mathematics education?

No. I just don't believe that the data support a conclusion that our schools have failed. The fact is the business community does not complain about the quality of the engineers, physicists, and chemists. However, I do think that we are far too complacent about the large proportion of our children who live in poverty, about the vast differences in educational resources between rich and poor school districts, about the rising costs of higher education, about reductions in the real value of student financial aid for low-income students, and about decreasing state expenditures for higher

education--and what that does to student motivation. Unfortunately, we assume that schools can be improved with little attention to the underlying conditions of poverty and often hold schools accountable for "fixing" the problems of society. My concern is that a focus on test scores deflects attention from what we can do to solve our real problems.

7. You talk about financial resources--isn't that throwing money at the problem?

Not at all. We know that low-income and minority students, on average, have less opportunity to study science and mathematics than do other students. They also have less access to the most qualified teachers and less adequate facilities and equipment for learning science and mathematics. We also know that there are large differences in education spending across rich and poor school districts. New York City spends \$7,299 per student. Great Neck, a wealthy suburb not far away, spends \$15,594 per student. If the amount of money spent on schools really doesn't make a difference, affluent parents haven't yet heard the message. If Manhattan had as much funding as Great Neck, it could make significant reductions in class size, increase teachers' salaries, buy computers for all the classrooms, and still have enough left over to finance decent science laboratories.

8. Don't you believe that national testing would improve education for the students you are most concerned about?

To the contrary, I am concerned that there would be serious, negative consequences. Harold Howe II, a former Commissioner of Education, has described the potential impact of testing on students from low-income backgrounds who have major problems to overcome both outside and inside their schools. He puts it this way: "Inside their schools, they are subjected to the effects of lower educational expenditures per student--larger classes, limited special services, decaying and inadequate facilities, higher levels of teacher turnover and teacher absence, and numerous other signals that they are second-class citizens of the education system. To remind them with a new national test of these discouraging facts is not the best route to building their morale or their performance." And, of course, for all students there is the problem I referred to earlier about increasing the emphasis on multiple choice tests of highly selected pieces of information. However, it is often argued that the way to implement an effective national testing system is to develop innovative new tests--performance assessments, essay exams, portfolio assessments. In how many years? At what cost? In recent testimony before the House Committee on Education and Labor, Dan Koretz, George Madaus, Ed Haertel, and Al Beaton estimated the cost of administering tests nationally in five subject matters in only three grades at more than \$3 billion per year. It simply means more money for consultants and less for students! By comparison, the entire Chapter I program, the largest

federal program for elementary and secondary education, spends \$6.7 billion.

9. What about American competitiveness? How can we compete in the global marketplace with the Japanese (the Koreans...the Germans...) if our students don't do better in these test comparisons?

Our problems in international competitiveness do not relate to weakness in science education or international test comparisons, but to business practices, government policies, and the realities of a global economy, for example, the lack of incentives for industry to invest in long-term product development, the financial incentives that lead to off-shore manufacturing, differential wage rates among countries, differential government subsidies among countries, licensing practices, antitrust concerns, and the emphasis placed on military at the expense of civilian research. The U.S. spends a smaller proportion of its resources on civilian research and development than do Japan and Germany. Approximately one-third of total U.S. expenditures (and two-thirds of federal expenditures) for research and development go to defense. That resource allocation hurts the competitiveness of the private sector to the extent that the resources could have been used to support commercial research and development leading to marketable products. These are far more important explanations of the status of U.S. competitiveness than are rankings on international test comparisons.

10. And a concluding question: Will we be first in the world in science and mathematics by the year 2000?

It all depends on the measure. If we choose our sample carefully, for example, only from the students attending the Bronx High School of Science, we will be first! Or, more seriously, if we can somehow measure our students' expertise in designing independent research projects, as demonstrated by the Westinghouse Science Talent Search, we will do quite well. The fact is our basic scientific research output is highly competitive right now--Nobel prizes, scientific publications, high quality scientists and engineers. However, if we continue to measure ourselves by international test scores, we will be far from first place. Far more important are other measures--those I just mentioned and others such as the vitality of our labor force, the employability and wages of those who do not attend college, and the quality of our workplace training. Yes, our schools can be strengthened, but our success in doing so will require us to focus on a number of difficult public policy issues rather than on test scores and rankings that tell us very little about how to resolve or even identify the most serious problems.