

NOTES FOR DISCUSSION OF INTERNATIONAL COMPARISONS OF  
SCIENCE AND MATHEMATICS EDUCATION

Comparative Education Class  
Teachers College, Columbia University

April 29, 1992

Iris C. Rotberg

Thank you for inviting me this afternoon. My articles on international comparisons have made two main points: (1) international test score comparisons are flawed methodologically and do not reflect the quality of education in any country, and (2) a reliance on a narrow criterion--answers on multiple choice tests--ignores far more important measures of our strengths and shortcomings in science education and leads to trivial, and sometimes counterproductive, solutions. I would like to organize today's discussion around a set of questions that I am often asked about these points. 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, the students represented in the test comparisons are much more highly selected in some countries than in others:

o High school attendance rates higher in U.S. than in most other countries. U.S.--80%; Europe--average 20% (in 1960s).

o Major reversals of rankings between higher and lower grades--Hungary and England/Wales; Japan and Hong Kong; the former Soviet Union, Slovenia, and the United States.

o The point is that the more students who take the test, the lower will be the average score. That score has little to do with the quality of education in any country.

o Analogous to SAT scores. Ranking of states. Letter from resident of affluent and therefore high SAT district that recently merged with low-income, low SAT district.

Second, some countries, like the United States, have a relatively high proportion of low-income students who are in school and taking the test. Relevant factor--gap, not absolute income.

Third, there are differences in curriculum emphases among nations that contribute to the relative rankings:

- o The subject matter taught, sequencing of curriculum, item selection, and real and intended curriculum.
- o Calculus example, including greater likelihood of taking calculus in college.
- o The way the stuff is taught.

Bette Bao Lord (the author of Spring Moon) puts it this way:

"As a fifth grader in Brooklyn's P.S. 8 . . . even before I had mastered fifty words of Brooklynese my teacher, Mrs. Rappaport, began asking me for my opinion on every matter that reared its hair, much less its head, in class. . . . I was flabbergasted by the fact that an adult--and not just any adult; on the contrary, my most honorable teacher--would solicit the opinion of a child--not just any child; on the contrary, an eight-year-old immigrant just off the boat. . . . And before long I came to realize that the merits of one's opinions were not the crucial point of the exercise. The crucial point was to air whatever opinions one had, and today I value this aspect of what we Americans delight in praising as our way of life perhaps more than any other. To me, the cacaphony of puddingheads offering their views is preferable to the clarion call of even the greatest emperor."

You might conclude that I believe U.S. education needs no improvement. I know we all wish that were the case. I don't think any of us would disagree that our education system can be better than it is. I am concerned, however, that we are in danger of losing some of our current strengths by pursuing an elusive gain on standardized test scores and that in the process we will ignore far more important problems.

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 participating countries.

Second, major inconsistencies in test score results that could not be explained by differences in the quality of education:

- o Major reversals of rankings between higher and lower grades, described earlier.
- o Contradictory findings about how well high scorers do in the United States compared to other countries.

o Contradictory explanations of test score differences, for example, between the United States and Japan--Harold Stevenson (teaching for understanding); Ian Westbury (curriculum differences); Karel Von Wolferen, author of The Enigma of Japanese Power (memorization).

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. Problems endemic to all studies for past 30 years.

o Not a matter of statistical expertise, but of societal and educational diversity among countries.

o Also large differences among countries in which students take the test:

--Exclusion of 20% of the classes.

--Exclusion of apprenticeship programs.

--Tracking by age 11.

--Highly specialized curriculum (even Princess Diana (not then a Princess, of course) did not continue past age 16).

--Problems magnified by inclusion of broader range/developing countries--most children will not be tested; out of school; highly elitist school systems; logistical problems; language problems; political problems. Tests are no more useful to a developing country than they are to the United States.

The difficulty of addressing these problems is illustrated by the most recent ETS study which went out of its way to point out the problems and strongly advised in its press release against ranking the countries. (See attached list.)

Inadvisable to take "heroic" measures. 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. Do you believe that we can learn something from other nations' education systems or teaching practices?

Of course we can. The challenge is to identify those practices that can realistically be transferred from one nation to another.

o However, in most cases, it would involve a basic restructuring of a nation's social, cultural, and political structure, including changes in the respective roles of national and local governments in education, the role of the teacher in society, teachers' salaries, comprehensive high schools, 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.

o Even when there is a public discussion within a country about making basic changes in education, the nation's social and cultural structures make it very difficult to accomplish. For example, in the United States, the current debate is about giving our elementary and high school students a more demanding curriculum, and then testing them on it; Japan would like its students to express their own views more readily; Taiwan would like its students to play more! A matter of culture, not the education system. In the area of international competitiveness, there are similar problems in trying to adapt industrial policies from Japan or Germany to the United States because government/industry links differ so fundamentally between the countries.

o R.A. Garden, Director of Research and Statistics, Department of Education, New Zealand, and IEA General Assembly representative, described the difficulties in applying "principles" from other countries this way:

"What appears to be a growth industry for U.S. educational researchers, comparing the USA and Japan, is likely to be fruitless [in terms of suggesting policy reforms]. Factors which influence opportunity and motivation to learn are too different in these widely differing cultures." (from Seth Spaulding article)

5. 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, Ph.D. physicists who can't find jobs (813 applicants for one position at Amherst College) and engineers who can't find jobs, despite claims of shortages.

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, we are far too complacent about:

o The large proportion of our children who live in poverty--the

numbers of poor children have increased in the 1980s. (See attached article.)

- o The vast difference in educational resources between rich and poor school districts.

- o The rising costs of higher education, reductions in the real value of student financial aid for low-income students, and decreasing state expenditures for higher education--and what that does to student motivation.

My concern is that a focus on test scores deflects attention from what we can do to solve our real problems. The current rhetoric assumes that schools can be improved with little attention to the underlying conditions of poverty and often holds schools accountable for "fixing" the problems of society.

Expand on school finance and higher education (access, costs, student financial aid and loans, polarization, teaching/research).

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

Not at all:

- o Low-income and minority students have fewer opportunities to study science and mathematics. Example of superintendent who moved from the West Side of New York to Harlem.

- o Large differences in education spending between rich and poor school districts (New York City--\$7,299 per student, Great Neck--\$15, 594 per student), and how the money could be used.

--If the amount of money spent on schools really doesn't make a difference, affluent parents haven't yet heard the message. Or, as a judge in a school finance case concluded:

"If money is inadequate to improve education, the residents of poor districts should at least have an equal opportunity to be disappointed by its failure." (as reported by Jonathan Kozol)

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

To the contrary, it is likely to have serious, negative consequences:

- o Problem of raising course and graduation requirements without doing anything about the vast financial differences between rich and poor school districts. These requirements will do more harm than good by reducing graduation rates and subsequent employability and earnings, increasing tracking, and generally screening out of

the education system those students who already receive the lowest quality education. 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? Harold Howe II, a former Commissioner of Education, describes the potential impact of national testing on students from low-income backgrounds, who have major problems to overcome both outside and inside their schools:

"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."

o Problem for all students of placing increased emphasis on rote learning, as measured by multiple choice tests, and less emphasis on basic scientific concepts, research methods, independent research projects, and public policy issues. (See attached letter to the editor.)

o Problem of trying to develop "innovative" new tests--performance assessments, essay exams, portfolio assessments. In how many years? At what cost? Estimate of more than \$3 billion per year for five subject matters in only three grades (testimony before the House Committee on Education and Labor by Dan Koretz, George Madaus, Ed Haertel, and Al Beaton). More money for consultants, less for students! Better to spend it on Chapter I.

o Multiple, and often contradictory, uses of tests.

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:

o Examples are the lack of incentives for industry to invest in long-term product development, 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.

o The United States 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:

o Choose sample from the Bronx High School of Science!

o More seriously, if we can measure our students' expertise in designing independent research projects, as demonstrated by the Westinghouse Science Talent Search, we will do quite well.

o The fact is our basic scientific research output is highly competitive--Nobel prizes, scientific publications, high quality scientists and engineers--and no shortages.

o Fewer U.S. students are going into Ph.D. science programs, but that trend results from a range of factors unrelated to test scores--the job market in other fields, the length of graduate study, and the need to pay off student loans. But perhaps these trends will begin to turn around as the job market in other fields, for example, investment banking and law, declines.

o And, of course, the United States has always been fortunate to attract foreign students and scientists. Einstein and Fermi and a host of other gifted immigrants have contributed to our scientific productivity.

o If we continue to measure ourselves by international test scores, we will be far from first place.

o Far more important are other measures--those I just mentioned and others--the vitality of our labor force, the employability and wages of those who do not attend college, the quality of our workplace training, and the extent to which minorities and women have the opportunity to enter science and engineering fields.

o 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.