ISSUES IN SCIENCE AND MATHEMATICS EDUCATION

Summary of Remarks at the Spring Meeting
of the Southern Technology Council
May 22, 1991

Iris C. Rotberg*

I think it would be useful to describe what I believe to be the real problems of U.S. education and to distinguish them from those problems that are highly publicized well beyond their significance. I will do so by using as a vehicle to point up real vs. unreal problems the highly publicized comparisons between the performance of U.S. students and those in other countries. These studies are reported to show a great inferiority of the U.S. education system. That conclusion, which I believe is inaccurate, has deflected our attention from painful and difficult problems. So too has the rhetoric over U.S. competitiveness which is assumed to be a deficiency attributed primarily to the quality of U.S. science, engineering, and technology education. Also a myth is the conclusion that the problems of education can be resolved without attention to the underlying conditions of poverty. I will also discuss other myths about U.S. educational practice and how focusing on them, while rhetorically satisfying, not only does little good but inevitably causes us to recommend solutions that are irrelevant at best but typically are counterproductive to resolving or even addressing real problems. For example, developing tough national standards on tests that many low-income students will inevitably fail will not increase their educational achievement, prepare them for a productive life, or lead them to skills necessary to compete in an increasingly competitive and demanding workplace. But I am getting ahead of myself. Let's talk a bit first about all those international comparisons.

As you know, there has been a lot of dumping on the U.S. education system. Typically, the evidence used to show that our schools have failed are the results of the international comparisons of science and mathematics test scores.

Ever since these comparisons began in the 1960's, Americans have believed the myth that U.S. students are outclassed by those in other nations. Yet, after almost three decades of apparent failures on international tests, we have somehow managed to maintain a level of research productivity that is overwhelming.

---

*Iris C. Rotberg is a Program Director with the Directorate for Education and Human Resources, National Science Foundation. These remarks are based in part on more extensive analyses that appeared in the December, 1990, and June, 1991, issues of Phi Delta Kappan. The views expressed are those of the author and do not necessarily reflect the positions and policies of the Foundation.

The apparent contradiction between low test scores and high research output would be only of casual interest if the test score comparisons were not taken so seriously. I am concerned that the public policy dialogue will continue to focus on test scores rather than on the far more important questions about our accomplishments --and our problems--in science and engineering education.

The fact is international comparisons of test scores are highly misleading indicators of the quality of our education system or of the expertise of our students.

In my remarks today, I make two main points:

First, the rankings among nations in international test comparisons are meaningless because the studies are unsound--the major educational and societal differences among nations make it virtually impossible to conduct valid comparisons.

Second, I argue that a criterion of test scores as the primary indicator of achievement in science and mathematics--even if it were accurate--does not reflect our accomplishments or, more important, address our real problems. Indeed, a preoccupation with test comparisons deflects public policy away from the important problems and I believe will lead us to implement "solutions" that are counterproductive to the long-term improvement of science education.

To my first point:

The international test scores are based on biased samples:

High school attendance rates are much higher in the U.S. than in most other countries. The original assessments compared the average score of more than three-fourths of the age group in the U.S. with the average score of the top 9% in West Germany, the top 13% in the Netherlands, and the top 45% in Sweden. It's not surprising U.S. students did not do well!

Of course, this type of problem is not limited to international test comparisons. The relative rankings of states on average SAT scores are also a reflection of the proportion of students who take the test. The states with the highest proportions of students taking the SAT tend to have the lowest average SAT scores. The best way to increase a state's ranking would be to discourage students from applying to college!

Recent international assessments have tried to deal with the sampling bias by testing only those 12th grade students who are still taking mathematics, but it has not solved the problem. Consider, for example, Hungary and England/Wales. Hungary ranked near the top in the eighth-grade comparisons; not surprisingly, by the 12th grade, when Hungary retains more students in mathematics than any other country, Hungary ranks among the lowest countries. Have Hungary's schools gone down hill between the eighth and the

12th grades or is it simply a matter of more students, lower scores? England/Wales, in contrast, scores among the bottom countries in the eighth grade comparisons but moves quickly to the top by the 12th grade when only a highly select group of students is taking the test.

Similarly, in the eighth grade comparisons, Japan ranks first, with Hong Kong in the middle of the distribution. By the 12th grade, when only 3% of Hong Kong's young people are taking mathematics (compared to 12% in Japan), Hong Kong comes in first and Japan second. The reality is that Hong Kong's schools are not dramatically better in the 12th grade than in the eighth; the changed rank is simply a matter of extreme student selectivity in Hong Kong.

The point is 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.

While almost everyone agrees that previous studies have been flawed, the argument is made that we will do better in the future. Yet for close to 30 years we have had expert statisticians working on the problems. They have not been able to solve them. Why? Not because they were unable to develop elegant statistical designs but because it was unrealistic to attempt to implement these designs in the real world. It is simply not a statistical problem. It is a problem of trying to compare highly diverse societies and education systems.

Each country differs in its division of students by language, social class, ethnicity, race, religion, immigration status, region, public or private schools, academic or vocational schools. We can't even describe accurately the countries' educational systems much less compare them.

In some countries, significant numbers of low-achieving schools--or schools in which the curriculum is considered to be inadequate--are excluded from the comparisons.

In other countries, many students who are in industrial apprenticeship programs do not participate in the test comparisons.

Several countries track students for all subjects in separate classrooms or separate schools as early as 10 years of age. A student's track is highly correlated with social class. We don't know which students are represented in the test comparisons and therefore we can't determine the reasons for a particular level of test performance.

In some highly elitist education systems, students take courses primarily in their field of specialization after age 16. Therefore, high school students who are tested in science and mathematics have studied essentially <u>only</u> science and mathematics from age 16 on. These students are compared with students in

comprehensive schools in other nations.

The problems are magnified enormously by the inclusion of a much greater range of countries in the forthcoming studies. The Soviet Union has been added to the list as well as Brazil (with 39% of the 12-17 year olds in school), China (with 43% in school), and Mozambique (with only 5% in school). China is an interesting case. Because of scarce resources, China has a highly elitist school system that provides a high quality education to a very few selected students who are in key schools. The vast majority of the students have not covered the material in the assessments and in fact will never be tested. The problem is compounded by great differences between urban and rural areas. A sampling of elite schools in China would distort the results in the same way as a U.S. sample composed primarily of students from the Bronx High School of Science. The problems are similar in other developing countries.

It is then argued that the solution is to sample the entire country. It won't happen. It can't be done. Most children are out of school by the age the tests are given. In Brazil, millions of children are homeless. In Mozambique, only 5% of the age group is in secondary school at all. There are political realities in testing children in regions that have a tenuous relationship with the central government. There are logistical problems in reaching remote rural areas. Typically, only the elite language groups can be tested in each country--Mandarin-speaking children in China; Russian-speaking children in the Soviet Union. The result is that a very small group of students--the highest achieving students--are tested in these countries compared to a much broader cross-section of students in the United States.

In short, these tests are no more useful to a developing country struggling to maintain an appropriate balance in the allocation of its scarce educational resources than they are to us.

Some have responded by saying let's take heroic measures, let's make sure there are strict international controls on who is tested, let's make sure we test out-of-school children (and by implication, I suppose, homeless children!). My response is why would we <u>want</u> to do it--how would these extreme measures (even if they were effective) improve the education of children in any of the participating countries?

I am sometimes asked whether I believe that <u>sampling</u> flaws are the only reason the U.S. scores relatively low in the rankings. Of course not. However, just because sampling alone probably can't account for all the differences, it doesn't mean the quality of U.S. education <u>does</u> explain the results.

For example, an important reality is that the U.S. has a higher proportion of students in poverty than many other industrialized nations--an unfortunate fact of U.S. society having nothing to do with the quality of education. Of course, developing countries

have the highest proportions of children in poverty, but many of these children are out of school and therefore not tested. There is an unrealistic expectation that the education system by itself--without fundamental changes in the underlying conditions of poverty--can deal with the educational problems associated with poverty.

Curriculum differences from nation to nation also affect the test results. For example, advanced mathematics students in the U.S. are more likely to defer calculus until college than are their counterparts in many other countries.

While there is room for debate about whether a higher proportion of U.S. high school students should take calculus, this issue cannot be resolved by examining the results of international comparisons. If we think it wise to teach calculus to a larger proportion of twelfth graders, let's do so after an analysis of the issue on its merits--Who would teach it? What course would it displace? Are students who take calculus for the first time in college at a disadvantage?--and not on the basis of the test scores of students who have never taken the subject.

Bette Bao Lord makes another point--the <u>way</u> stuff is taught in the U.S. She 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." (Bette Bao Lord, <u>Legacies A Chinese Mosaic</u>, New York: Alfred A. Knopf, 1990, pp. 99-100)

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 <u>losing</u> 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.

That gets me to my second point.

Let us <u>assume</u> that test results accurately portray the relative "rankings" of participating countries. We are still left with the

matter of whether these scores are useful measures of those things that are most important to us--or to other nations--in science and engineering education. I would suggest that even a methodologically sound study of test performance does not address far more important issues with respect to science and engineering. The comparisons clearly do not reflect the breadth of a nation's accomplishments or address its real problems. For example:

How productive is the U.S. in basic and applied research fields? What does the marketplace say about the research opportunities in our institutions of higher learning? Where are students from other parts of the world taking their advanced degrees in science and engineering?

What are our accomplishments in making major technological advances, as measured by patents and their use in products? Are we successful in turning our scientific and technological advances into products that are competitive in the international marketplace? If not, why not? Does it have much to do with our educational practices?

Are the fields of science and engineering attracting high-achieving students? Is there a shortage of students or faculty members in these fields? Are we making progress in attracting women and minorities to these fields?

Does the teaching environment give students who do not major in science and engineering some understanding of key scientific issues and methods?

Are we providing the general student population with an opportunity to gain the skills that are needed in order to be competitive and productive in the workplace? Are we maintaining the technical expertise of the workforce?

The answers to these questions are mixed, but they are far more meaningful measures of our national accomplishments and problems than are comparisons of test scores.

A few examples may be useful here:

There is little question that our research productivity is extremely high. The U.S. accounts for more than a third of the world's scientific publications. The next highest-ranking competitors are the United Kingdom, Japan, and the Soviet Union, at about 8% each.

It is also generally acknowledged that no other nation's system of higher education offers the breadth and quality of research opportunities available to students in U.S. universities.

Our international competitiveness, of course, is not good in some areas, although there are recent examples of increased exports because of greater efficiency and quality control and because of a

cheaper dollar. While our competitiveness clearly depends on a highly skilled labor force, it also reflects a number of other factors: the global economy, the lack of incentives for industry to invest in long-term product development, business practices with respect to off-shore manufacturing, interactions between universities and industry, the emphasis on military at the expense of civilian research, the lack of a civilian technology policy. Approximately one-third of our total R&D expenditures--and two-thirds of federal R&D expenditures--go to defense.

It is not surprising that technologies in which the U.S. is particularly competitive are technologies--for example, biotechnology--that are close to basic research without the intervening steps of lengthy technology development or heavy capital investment.

There has also been a lot of rhetoric about shortages of scientists and engineers. While there may be shortages in some regions or spot shortages in some fields, the fact is that there is little evidence of serious shortages nationwide. Indeed, in some fields it is difficult to find a job--there are too many, not too few applicants.

Bachelor's degrees in engineering awarded to U.S. citizens and permanent residents almost doubled between 1977 and 1987. While there have been some declines in the numbers of students choosing the physical sciences and mathematics, and some shortages of precollege and college faculty in certain fields and in certain regions of the country, again, these shortages have little to do with the quality of education. They are related instead to other considerations--salaries, working conditions, competition from other fields like business, investment banking, and law.

The fact is that there is no problem with the supply of highly qualified students. The skills of our science and engineering majors are extremely high--for example, students who choose to enter science and engineering fields continue to rank well above the national average on academic measures. And an analysis of SAT mathematics scores shows that these scores have actually risen in recent years. In 1977 the 90th percentile score was 628; in 1986 it had risen to 642.

An important question is whether our science and mathematics courses focus on the important stuff--major scientific concepts, scientific issues in the context of public policy, some understanding of what the scientific method is all about, and in the case of mathematics, statistics and problem solving. Unfortunately, the increasing emphasis on standardized tests, even apart from the international assessments, encourages the teaching of large amounts of superficial facts to be memorized and makes it more difficult to implement the type of curricula that are generally acknowledged to be desirable.

Finally, our most important problem--and our most difficult

challenge in science and engineering--is providing a better education for the general student body in a world requiring ever-greater technological skills. I know all of us share a concern that U.S. society will grow increasingly polarized if a significant proportion of our population lacks the skills needed to compete for jobs that provide a reasonable income. And because poverty correlates so highly with educational problems, these problems are likely to be exacerbated over the years if the current trends continue.

These are difficult problems, but they are unlikely to be alleviated by yet another round of tests. In fact, a preoccupation with tests may lead us to "solutions" that are at best trivial and may be counterproductive to dealing with the tough questions.

The current rhetoric assumes, for example, 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.

The current rhetoric assumes that we can "fix" our schools by giving more tests.

It assumes that low-achieving children will do better if they are given yet further evidence of their failures--low test scores.

It assumes that teachers aren't trying hard now, but that they will try harder if only we administer more tests.

The current rhetoric ignores the potential negative effects of a national test (or something close to it) on low-income students and the likelihood of increased failures, increased tracking, lower graduation rates, fewer subsequent job opportunities.

Harold Howe II, a former Commissioner of Education, puts it this way:

"One way to teach kids to swim is to throw them into deep water. Some will swim, and some will drown. A better way to teach kids to swim is to give them some swimming lessons in shallow water, and then to give them the chance to demonstrate their skills in water that is over their heads. . . . My notion of the place to start new learning and assessment modes is in the classroom-not with a national test two or three years from now. A predictable result of springing such a test on teachers and students will be a strong message of their inadequacy. What good will that accomplish? We already know it from NAEP.

"Within the framework that characterizes these ideas for change in school practice is a very positive theme about the motivation of students. This theme grows from opportunities for more exploratory than rote learning, for students doing more talking and less listening in their classrooms, for working in groups to fashion evidence of their learning, for dealing with open-ended questions,

for students suggesting questions that might become part of their school studies, for deeper understanding of the world around them, and for becoming thinkers about that world rather than confused participants in the struggle for competitive survival. It is a theme that holds the potential for maintaining the curiosity of students far more powerfully than evidence of failure on a test.

"Those students from disadvantaged backgrounds will, I think, respond to such motivating classroom practices. But we have to remember that they contend at the same time with two major handicaps-one of them inside their schools, and the other outside them.

"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. . . . " (Harold Howe II, letter to Jack Jennings, April 29, 1991, pp. 4-5)

Which brings me to one more assumption of the current rhetoric--money doesn't make a difference.

We all know that there are still enormous differences in education spending across rich and poor school districts. A recent New York Times article reporting on a school finance case in New York state describes the contrast this way:

"Just a few miles from this Suffolk County community, there are public school districts that spent more than $17,000 for each student last year, with schools that boast of computers in every classroom and a vast array of extracurricular programs.

"But in the William Floyd School District here, which spent $7,614 for each student, nearly half the students attend school in temporary trailer-style buildings, because there is no money to build new classrooms. At the high school, built for 1,300 students, the enrollment is 2,000, and lunch starts at 9:06 A.M.--that is the only way the school can avoid violating fire-safety codes in the cafeteria." (Sam Howe Verhovek, "Poorer New York School Districts Challenging State Aid as Unequal," The New York Times, May 6, 1991, p. A1)

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 gotten the message.

There is clear evidence, in addition, that low-income and minority students, on average, have less opportunity to study science and mathematics than do other students. They have less access to the

most qualified teachers, to adequate facilities and equipment for learning science and mathematics, and to the types of curricula and instructional strategies considered particularly effective with all students.

Although there has been more attention to the problems of rising costs of higher education and declines in the real value of student financial aid than to the other issues I have mentioned, the implications for low-income and minority students have not been fully addressed. Some students may choose not to attend college at all. Others may choose low-cost institutions, which has led to the increasing polarization of higher education. Still others may not be able to transfer from 2-year to 4-year colleges because of personal finances or because the state has had to reduce spaces due to budget cuts. Of course, decisions about graduate school and field of study are also affected. Not surprisingly, recent research has shown the close relationship between the availability of financial aid and minority students' decisions about whether or not to attend college. And because educational attainment is highly correlated with the level of parental education, our educational losses are not simply "one-time" misfortunes for this generation of young people--or for the society at large. They have enduring implications for future generations.

In short, even if our international studies were sound, how would all these tests improve the education of children in any country? There are a great many important questions related to the state of science and mathematics education in the U.S. and to the matter of <u>which</u> students need to be better served. Let's get on with it and focus our attention on the difficult public policy issues to be addressed rather than on comparisons and rankings.