all the available funds, would inevitably prevent development in these critical areas.

From Lawler's article (p. 282), it appears that ITER finds its strongest support in a "wealthy and influential association of major corporations. . . ." This sounds like an ominous repetition of history, as our problems today with nuclear fission power plants originated when the nuclear industry decided to bring to prominence the first fission reactor concept that appeared to work. Similarly, the adoption of this probably faulty device would have catastrophic consequences for the development of nuclear fusion energy.

Ernesto Mazzucato
Princeton Plasma Physics Laboratory,
Princeton University,
Princeton, NJ 08543, USA
E-mail: mazzucato@pppl.gov

Several recent letters proclaim once again the superior promise that thermonuclear fusion offers for future large-scale generation of electric power (D. E. Baldwin and T. C. Simonen, 13 Oct., p. 220; E. M. Campbell and J. C. Browne, 12 Jan., p. 130; and N. A. Davies, 12 Jan., p. 133). They point out the significant progress made, the benefit that the so-called "multiple mission" fusion offers, and the peripheral advantages that would accrue. While somewhat exaggerated, these arguments are correct, but they miss the point (see W. E. Parkins, Letters, 24 Nov., p. 1281). Unless the primary mission is fulfilled, all of the touted side benefits are but academic.

Can the fusion reactor concept *ever* be a practical, cost-effective method for production of central station electricity? Unfortunately, nature has interposed not one, but three, unsurmountable obstacles. (i) The required temperature and other plasma conditions for even the easiest fusion reactions (deuterium-tritium or D-T and the deuterium-deuterium or D-D) seem to be unattainable. (ii) These reactions release energetic damaging neutrons that change the physical properties of the reaction vessel and make it radioactive. (iii) Most devastating of all, power cannot be extracted from within the reacting plasma. It can be gathered only at the peripheral wall.

Each of these obstacles bears on the practicality and cost-effectiveness, and thereby the future, of fusion power. Efforts to achieve the necessary plasma conditions are leading to a reactor system design of monumental complexity. Effects of the neutrons dictate that the operating utility be prepared to periodically replace the highly radioactive and almost inaccessible vacuum vessel—an unacceptable requirement. And inability to extract power as heat within the reacting region (as is done in fossil-fueled boilers and fission reactors) forces the engineering to physical dimensions that are much too costly. When the utility executive finally figures the capital investment charges and operating costs that would apply to each kilowatt-hour generated, he will close the book on another broken dream.

But all need not be lost. There is no lack of scientific and technical frontiers, and there will be new ideas. The greater benefit will come from applying our resources, including the efforts of talented scientists and engineers, in directions that can make a difference in the future. When large-scale usage is intended, however, one must be sure that the development is guided by a practical and cost-effective concept.

William E. Parkins*
20120 Wells Drive,
Woodland Hills, CA 91364, USA

*Former Director of Research and Technology, Energy Systems Group, Rockwell International

■

Comparing Student Test Scores

In the Policy Forum "Myths about test score comparisons" (1 Dec., p. 1446), Iris C. Rot-

berg skillfully describes difficulties in conducting and interpreting tests of what children learn in school. The five myths are indeed myths. Yet her associated argument against the credibility of properly adjusted test scores is logically troubling. Unadjusted test scores, she correctly states, are grossly misleading for comparing the quality of schools because they are affected by many other factors as well. Then she invites a "catch-22" by also stating that statistical adjustment to filter out the influence of such factors "has not worked." Rotberg's alternative is to judge an educational practice "on a careful consideration of the merits." She does not say, however, how merit should be determined. Surely, judgment of merit would require evidence of student learning. Otherwise, we are thrown back on the myriad we-know-a-good-school-when-we-see-it contradictions of expert opinions and educational philosophies. Evidence for merit in science education at one time was children seated quietly in rows with work sheets; currently, it is noisy children at tables making up their own hands-on investigations; tomorrow, it may be children sitting slack-jawed in front of computer terminals. Education is as much a victim of untested fads as it is of fads in testing.

Credible comparisons of test scores do seem possible because Rotberg finds them good enough to prove the impact of other factors: "Research shows, however, that per pupil expenditure, teacher expertise, and class size do make a difference in student achievement." (She also says that tests can be useful to describe national trends.) A useful follow-up Policy Forum would discuss how those differences could have been demonstrated in spite of the alleged impossibility of comparing the achievement of different groups of students through testing.

Andrew Ahlgren*
University of Minnesota, Minneapolis, MN 55455, USA

*Professor Emeritus of education

Response: Student scores on standardized tests—whether or not they are statistically adjusted—are highly misleading indicators of the quality of schools and, therefore, do not provide a valid basis for an accountability system. Most of the problems involved in test score comparisons cannot be handled simply by making statistical adjustments, which do not capture key factors that invalidate the comparisons—student mobility and turnover within a given school year, incentives to encourage certain students to take or not take the test, or the consistency between the test and the instructional program. It is difficult, therefore, to interpret test score fluctuations or attribute them to changes in school quality (*1*). For example, a study of Title I, the federal education program for disadvantaged children, found that about one-half of the schools identified as needing "program improvement," based on test scores, appeared to be doing just fine only 1 year later—without making any changes in their Title I programs (*1*).

Even more disturbing are the negative consequences of test-based accountability systems. These systems provide incentives for schools to encourage low-achieving students to drop out before the test is administered. They also encourage the teaching of a narrow set of measurable skills that often have little to do with what educators and parents value most. The mandated tests—and the rote learning associated with them—are particularly common in classrooms with large proportions of low-income and minority students (*2*).

My suggestion that the decision to adopt a particular educational practice should be based on a careful consideration of the merits of the proposal (not on rankings on standardized tests) responded to the circular reasoning that occurs in discussions of in-

ternational studies. The argument holds that more high school students should study calculus because students who take calculus do better on the international mathematics tests, which include calculus problems. My point is that an analysis of such a proposal should be based on the merits—Are students who take calculus for the first time in college at a disadvantage? What courses would high school calculus displace? Who would teach it?—and not on the basis of the lower scores of students who have never taken the subject (3).

Finally, research on school expenditures showing, for example, the value of lower class size cannot help us design better test-based accountability systems. The reason is that the results of the former type of research do not have publicized negative consequences for teachers and students. As long as test score comparisons are used for accountability purposes, there will be an incentive to find a way to modify the test-taking population or "teach to the test" to achieve favorable results. Methodological and statistical analysis will not solve that real-world problem.

Iris C. Rotberg
7211 Brickyard Road,
Potomac, MD 20854, USA

References

1. M. A. Millsap *et al.*, *The Chapter 1 Implementation Study, Interim Report* (U.S. Department of Education, Washington, DC, 1992).
2. G. F. Madaus, M. West, M. C. Harmon, R. G. Lomax, K. A. Viator, *The Influence of Testing on Teaching Math and Science in Grades 4–12* (Center for the Study of Testing, Evaluation and Education Policy, Boston, MA, 1992).
3. I. C. Rotberg, *Phi Delta Kappan* **72**, 4 (1990).

■

HHMI Awards

The announcement that the Howard Hughes Medical Institute (HHMI) is awarding its Research Resources Grants to 30 U.S. medical schools (Jocelyn Kaiser, "Med schools receive Hughes windfall," News & Comment, 12 Jan., p. 138; correction, 2 Feb., p. 583) is applauded, but something is left out of the analysis. The 30 schools who have received funding are, almost without exception, the most successful and well-endowed schools in the country.

Why should HHMI have chosen to lavish its affections on schools such as Stanford University; the University of California, San Francisco; the University of California, San Diego; the University of California, Los Angeles; Harvard University; and Johns Hopkins University, when the stated intent of the opportunity was encouraging to smaller schools with emerging excellence ("The scientific reputation of the medical school will not be the primary criterion for awarding funds")?

If HHMI had wanted to have an impact on the research enterprise, it might have spent more of its funds on schools with an excellent, if small, research enterprise. Schools like the University of Nevada School of Medicine (which, ironically, was started by a financial contribution from the late Howard Hughes to the Nevada legislature in 1969) would have been able to improve all aspects of their operation with awards the size of those announced by HHMI, while schools the size of those receiving the largest of the HHMI awards will be able to improve only one or two programs and add a few faculty at best.

Could it be that it is time for HHMI to broaden its perspective?

Iain L. O. Buxton
Dean for Research,
University of Nevada School of Medicine,
Reno, NV 89557–0046, USA

Corrections and Clarifications

In the news article by Jon Cohen "AIDS trials take on peer review" (News & Comment, 5 Jan., p. 20), the table on page 21 should have listed the University of Miami instead of Miami University. The University of Southern California, omitted from the table, ranked 11th, with a score of 153. The University of California, San Francisco, had a score of 156.

"The Bell Curve: A statement" (Letters, 5 Jan., p. 13) was an edited version of a statement by the federal advisory group listed in the address at the end of the letter, the National Institutes of Health–Department of Energy Joint Working Group on the Ethical, Legal, and Social Implications of Human Genome Research (ELSI Working Group). The names of two members of that group appeared at the end of the letter for the purpose of correspondence. They were not the sole authors.

In the 24 Nov. Perspective "Ensemble activity and behavior: What's the code?" by S. A. Deadwyler and R. E. Hampson (p. 1316), the citation at the end of the figure legend (p. 1317) incorrectly stated that the data were taken from reference (5). That citation should have read "[Data taken from (24)]." Reference 24 is S. A. Deadwyler, T. Bunn, R. E. Hampson, *J. Neurosci.* **16**, 354 (1996).

Letters to the Editor

Letters may be submitted by e-mail (at science_letters@aaas.org), fax (202-289-7562), or regular mail (*Science*, 1333 H Street, NW, Washington, DC 20005, USA). Letters are not routinely acknowledged. Full addresses, signatures, and daytime phone numbers should be included. Letters should be brief (300 words or less) and may be edited for reasons of clarity or space. Letter writers are not consulted before publication.