

“Truth-in-Testing” Attracts Diverse Support

Even though claims of scoring errors, cultural bias, and cover-up seem to miss their mark

A state senator in New York, Donald Halperin, is unhappy about standardized educational tests. He is unhappy because a clerk in his private law office was prevented from attending a top school by a low score on the standardized law school entrance exam. He is unhappy because he himself had failed the bar exam the first time around, after overlooking a whole slew of questions on a single page.

As a result, Halperin cosponsored a pioneering bill that would open standardized admissions tests to public scrutiny. The bill, which was signed into law by New York Governor Hugh Carey on 14 July, may well have national consequences. A similar bill is scheduled for markup this month by an education subcommittee of the U.S. House of Representatives.

Each gives test takers the right to review their answer sheets, the original questions, and the correct answers, once a test has been scored. They also permit anyone to review the statistical underpinnings of an entire exam—data heretofore entrusted only to authors of tests and selected psychometricians.

Supporters of the New York law claim it will expose test questions that are culturally biased and reveal errors in scoring. The law covers the standardized exams commonly required for admission to college, graduate school, law school, business school, medical school, and in most cases, the professions themselves. Consequently, most of the academic and testing community is up in arms. It claims the legislation will have no discernible effect on the tests, other than raising student fees. Under the new law, tests will be used only one or two times before being released, so more money will have to be spent on developing new questions.

Supporters of the legislation are fueled by more than just resentment. For ideological reasons, they dislike the near-monopoly of the testing business by the Educational Testing Service (ETS), headquartered near Princeton, New Jersey. They dislike the secrecy with which this firm and others go about their business. And some of them dislike testing per se, viewing it more as a barrier to advance-

ment than as an opportunity for objective evaluation.

All of these arguments gain immeasurable strength from the simple fact that many of those who are speaking out against testing have been subjected to, and perhaps humbled by, the tests. ETS alone administers 8 million exams annually. An ETS official once stated proudly that “we are the nation’s gatekeepers.” While debate was raging in the New York State legislature, Governor Carey noted that a standardized exam had once kept one of his children on the wrong side of a gate.

The argument attracted “virtually every major parent, teacher, student, and civil rights group in New York State,” as a flyer of the New York Public Interest Research Group accurately suggests. Included were the National Association for the Advancement of Colored People, the Puerto Rican Legal Defense and Education Fund, and a northeast research group known as the National Consortium on Testing. The Parent-Teachers Association and the National Education Association, a teachers lobby, pitched in; instructors frequently oppose standardized testing as a constraint on their freedom to teach what they wish.

As a result of the efforts of these organizations, bills similar to that passed in New York State have been introduced in the legislatures of Maryland, Texas, and Ohio. California has already passed a weaker bill requiring only that copies of tests like those being administered be placed on file with the state department of education. Each of these efforts has in part been eclipsed by the New York law, because data on national tests that will become available there can be disseminated elsewhere.

Although most of the various admissions councils that administer standardized tests have promised to comply with the New York law, the authors of the medical school entrance exam, the Association of American Medical Colleges (AAMC), have threatened to stop giving their tests in New York after the law becomes effective 1 January 1980. James Erdmann, who directs the AAMC test program, says his group is pressing for an amendment that would exempt the

Medical College Admission Test (MCAT), on grounds that disclosure of test questions after a single use would soon exhaust all the questions that could be devised. The American Dental Association (ADA), which sponsors the dental school admission test, says the same thing. “We have a series of two- and three-dimensional objects which the test subject must relate to,” says James Graham of the ADA. “There are only a small number of ways to represent a box or a triangle on a drawing.”

Erdmann, of the AAMC, offers no speculation on the fate of the 5000 students who take the MCAT every year in New York. “That remains to be seen,” he says. “All I can say is that, given the current circumstance, this is the decision we had to come to at this point.”

Although the AAMC is screaming loudest, the law primarily will affect the testing program of ETS. Under contract, ETS writes and administers several hundred occupational and educational tests, including the well-known Scholastic Aptitude Test (taken by 1.4 million high school pupils in the last school year), the Graduate Record Examination (taken by 314,000 persons), the Graduate Management Admission Test (taken by 166,000), and the Law School Admission Test (taken by 128,000). ETS writes the bar exam in almost every state, including the test that Halperin failed.

For most of the supporters of legislation, ETS is the chief target. Since its founding as a nonprofit corporation by the Carnegie Foundation and several education associations in 1947, ETS has doubled its revenues roughly every 6 years, reaching \$80 million last year and earning the praise of *Forbes* magazine as “one of the hottest little growth companies in U.S. business.” Such high income and overwhelming success attract charges of monopoly and big-brotherism. Attention is drawn to the firm’s headquarters on 400 acres: the eight staff buildings, the president’s expensive home, the duck pond, and the \$3 million conference-center hotel with tennis courts and swimming pool. Deer actually go bounding by within sight of the offices. “For sheer elegance, ETS’s conference center would make the Shah of



ETS's headquarters, replete with duck pond and conference center, draws allegations of profiteering at students' expense.

Iran weep," says Chuck Stone, who worked at ETS as director of minority affairs. Allegations are made of profiteering at students' expense.

ETS officials, for the most part, dismiss the complaints and indicate bafflement at what the fuss is all about. They claim with justification that hard evidence of ETS's failings is sorely lacking. The firm is big, they say, because it was shrewd enough to fulfill the needs of disparate educational institutions at the right time; it dominates the field because it hires the top experts and does its job well. It is somewhat secretive about its work because, well, tests have to be secret. As for the magnificence of the headquarters, this visitor found it comfortable but not plush; it has roughly the charm of a college campus that admits no undergraduates. The 400 acres were purchased long ago, and most remain undeveloped. The conference center was indeed constructed with student fees, but it is virtually self-supporting now, from rent charged to consultant contracts and visiting associations.

The idea of thousands of kibitzers venting their rage at test questions and scoring does not beguile ETS officials. On the other hand, they seem to realize there is merit to more outside scrutiny, even if only to mute criticism that finds as its wellspring a vacuum of information. "After we lost in New York, we had some internal discussions, because we were portrayed to the public as being very negative," says a high-ranking ETS official. "It is now our desire not to be adversarial, although we still have some problems with the law as passed in New York and introduced in Washington [D.C.]."

Whatever their current stance on the political merits of supporting reform, ETS officials consistently argue that the usefulness and validity of ETS tests is assured. "There have been occasions

when people have misused test scores"—not tests—acknowledges Robert Solomon, the ETS executive vice president. "Part of the current brouhaha is caused by those who want to eliminate the abuses of testing, and by those who sincerely believe that the policy of sunshine is a good thing. But there are also those who want to eliminate testing from education for political reasons."

Some of the proponents of "truth-in-testing" actually do hope to diminish its importance in admissions. One such critic is Banesh Hoffman, a mathematician at Queens College in New York. Hoffman has written that multiple-choice exams of the variety prepared by ETS "are inherently superficial, using vagueness, imprecision, and worse as substitutes for the worthwhile difficulty their format cannot encompass. They sap the strength and vitality of a nation by rewarding conformity and quick-witted mediocrity while penalizing depth, subtlety, individuality, and creativity." More moderate critics merely say the limitations of individual tests are overlooked. ETS exams, for example, are highly imperfect predictors of performance—ranking about .35 on a scale of 0 to 1 with 1 as perfect prediction; grades are considered superior diviners. A divergence in scores of 60 points on a 600-point scale is considered statistically insignificant. ETS maintains these limitations are widely publicized and understood, but one commentator compares such warnings with "adjuring smokers that cigarettes can be injurious to your health."

Nothing in the legislation passed in New York and proposed in Washington will directly limit the emphasis that schools place on the tests, or add to the tests' validity, as even the supporters concede. Their intention instead is merely to reveal deficiencies through open distribution of test questions, an-

swers, and statistics, and then to let public and professional opinion take its course. The argument has the advantage of obviating a need to prove that deficiencies actually exist, and of appealing to anyone who thinks more openness is a good idea. Sam Gibbons (D-Fla.), the sponsor of national legislation similar to that enacted in California, explains his support by saying, "it is time that the freedom of information laws, truth-in-lending laws, and truth-in-packaging laws be applied to the field of educational and professional testing and licensing." Representative Ted Weiss (D-N.Y.), the sponsor of New York-styled national legislation, voices the same thought, and compares the industry unfavorably with utilities. Both perform a public function, he says, but only testing is exempt from public scrutiny.

To some, all these exertions over openness would seem unavailing without the presupposition that ETS suppresses information damaging to its tests. Ernest Kimmel, the head of test development there, denies it, saying that any data ETS gathers are available for the asking to any competent researcher (except the names of students, of course). Several independent psychometricians contacted by *Science* agreed. Lee Cronbach, a well-known expert in educational psychology at Stanford University and a member of the ability testing panel of the National Academy of Sciences, says he has had "very little difficulty in getting the information I need for my research, though of course I may be in a special position." Cecil Reynolds, an expert in test bias at the University of Nebraska and new publisher of the respected annual, *Mental Measurement Yearbook*, agrees: "I have contacted ETS for information, and if you know what you want, you can write and they will provide it. The difficulty, however, is in knowing

(Continued on page 1113)

fano's written statement to the news media.

Jacobson says the lack of fanfare was due to the chaos that apparently reigned at HEW during Califano's closing days. Others contend that the controversial nature of the report made a low-key delivery more appropriate.

The report notes, for instance, that iron is the most common nutrient deficiency in America. The meat industry is quick to point out, however, that the report failed to mention that red meat is the "best source of biologically active iron." Says Carlson: "Blanket recommendations that red meat consumption be reduced are not without risk to millions of Americans."

Pop Nutrition Books Face Legal Hurdles

Hard times may be on the horizon for the publishers of self-help books that dispense medical advice off grocery-store shelves. A couple in Pompano Beach, Florida, is suing the estate of the late nutritionist Adelle Davis and her publisher, Signet Books, over the death of their 2-month-old son. If they win, legal experts say that fewer "health" books will probably be published.

In her book, *Let's Have Healthy Children*, Davis advised that infants suffering from colic should be given large doses of potassium. She cited a study of 653 sick infants in which a treatment of 3000 milligrams of potassium chloride was successfully used to correct the condition. Susan Pitzer of Pompano Beach says she followed Davis's advice. In April 1978, she gave her son 2500 mg of potassium chloride one day and 500 mg the next. After the second dose, the infant had cardiac arrhythmia and was rushed to a local hospital. Four days later he died.

The Pitzers' attorney, Peter A. Portley, told *Science* that Davis was negligent for not including a warning with her advice—that potassium should not be prescribed for dehydrated patients. "Without that caution," says Portley, "she was in effect giving lethal advice."

The suit seeks more than \$1 million in damages and claims that Signet

was grossly negligent for not including a warning and verifying advice in the book. Portley says he hopes to pressure publishers of self-help books into editing more carefully. A trial date has not yet been set. An attorney for Signet says that Davis stated in the beginning of *Let's Have Healthy Children* that she didn't expect her ideas to take the place of a visit to a physician. He also says that a "medical expert" reviewed the manuscript before it was first printed in 1951, and that it was "updated" in 1959 and 1972. Signet also recalled the book last spring so doctors could again revise it.

The Pitzer lawsuit is not the first filed against Davis, who died of bone cancer in 1974. Portley says that in 1976 a Maine woman followed the advice in *Let's Have Healthy Children* and gave massive amounts of vitamin A to her infant. The child is now stunted. She recently won an out-of-court settlement from Davis's estate for \$150,000.

Dollar-a-Gallon Gas Fuels Metric Conversion

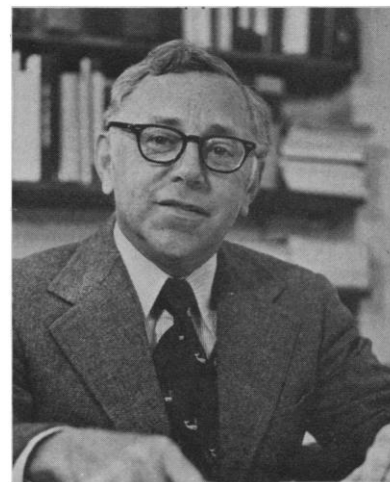
A glance at a gasoline pump shows that most pumps are equipped to record prices up to 99.9 cents per gallon. At the end of August, however, it was reported that the average price of gas in the United States for the first time crept over the \$1-a-gallon mark. The upshot is that many dealers are charging by the half-gallon, and then doubling the price shown on the dial. This, however, is not going down well with the federal weights and measures people.

One solution is for dealers to pay \$200 per pump and have an extra digit put on the price meter. But officials at the U.S. Metric Board say they have a better idea. For only \$50, a pump can be converted to measure out liters. In California, more than 50 service stations have already done so, and oil companies are running public-acceptance trials of the liter method in various parts of the country. Bill DeReuter of the U.S. Metric Board says that if all gas pumps in the United States were converted to metric instead of having an extra digit added, it would save the oil industry more than \$100 million.

William J. Broad

(Continued from page 1111)

precisely what they have to offer, as they don't routinely publish it." Both agreed that the law would force ETS to organize its data more carefully and publish a catalog for outside experts. It also would force ETS to release the raw scores of students, which it does not do now even on request. Raw scores are necessary to



Robert J. Solomon

verify the scaled scores that ETS does release, and to examine the computations for equating one test to another administered at a different time. "It is just our professional judgment not to let them out," says Kimmel. "It's not a cover-up; they would generate more confusion than they would clear up."

If pushed an inch, of course, most supporters of the testing legislation will freely admit to believing that a cover-up does exist. In the nature of all conspiracy theories, the sparseness of the evidence becomes an argument that suppression exists. ETS points out that an associate of consumer advocate Ralph Nader has been researching problems at ETS for 5 years, and reportedly intends to publish a book; as yet, however, no report has been issued, no scandal unearthed. Nader's people, on the other hand, suggest that with access to crucial data, horror stories would flow. An internal ETS report on cultural bias is an example of one such exhibit withheld; it is said to have found bias alive and well in ETS questions. ETS acknowledges that the report exists but says it is somewhat flawed and that a "complete extant copy" cannot be located at present. An ETS official agreed it would be useful to get the report out, dispelling incorrect assumptions about it; and suggested that *Science* could see a copy as soon as one is located.

Lacking this report, the supporters of legislation instead cite outside research

revealing a rough correspondence between family income and the scores on ETS tests; sure proof of cultural bias, they say. ETS's retort is simple: its tests only mirror inequities already present. "The Scholastic Aptitude Test takes society as it is, not as it ought to be," says Fred Hargadon, chairman of the College Entrance Examination Board, for which ETS prepares the test.

But what about scoring errors, ask the law's supporters; these, too, would be laid bare by the open return of questions and scored answers. Of two such errors

discovered, however, one was minor, and the other was technically a judgmental problem, not a mistake. The first occurred when 90 percent of those who took the Graduate Management Admission Test in the fall of 1977 had their scores recorded 10 points too low on a single section; it was caught by ETS itself, although an official admitted this was quite by chance. The second incident occurred at roughly the same time, and unraveled when admissions officers at private law schools noticed that perfect or near-perfect scores on the

LSAT were turning up with unusual frequency. It resulted from an independent ETS decision to create a greater point spread at the high side of the scale. Complaints were made, and the Law School Admission Council, which contracts with ETS, ordered it to revert to the old scoring method. Walter B. Raushenbush, president-elect of the council and professor at the University of Wisconsin law school, says that "ETS had not been sufficiently alert to the importance of test score comparability over time." This, and concern over "quality control, delays, and costs" have been factors in the council's recent decision to pare ETS participation in the law school admission process (the council has taken direct control of a grade-reporting system and the prelaw handbook, and is considering writing the tests itself).

Though such an example excites the law's proponents to considerable rhetorical flight, it is hardly enough to suggest incompetence or discrimination. ETS believes that without more substantial cause, there is no reason for legal intrusion into the testing business. "In our view, harmlessness is not a reason for passing a law," says Hargadon of the College Board. Alas, neither is it a good reason for not passing a law that a lot of people seem to desire, concedes Solomon of ETS. But there is a good reason, he adds. Distributing the exam questions after one or two uses will substantially raise the cost of writing new tests—forcing a rise in student fees of several dollars per test (they range now between \$5 and \$15), and perhaps the outright cancellation of special exams for small groups, such as military personnel, the handicapped, or those who observe the Sabbath on Saturday morning.

Idle threats, say supporters of the legislation, and this time, they appear to have ETS dead to rights. ETS profit from each student fee amounts to 22 cents on the dollar, officials concede. Test development amounts to 7 cents. The balance is spent on services—studies, conferences, and publications—provided gratis to member schools of the admissions councils that contract with ETS for exams. Solomon insists these services are "just as important as the tests themselves. Students are indirectly benefiting from this use of their fees." Doubtless many students would not agree.

It would seem, then, that both the arguments for and against the law lack a certain standard of proof. One thing seems certain: both Houses of Congress are unlikely to act until the consequences of the New York law are clearly within sight.—R. JEFFREY SMITH

Test Coaching Dispute Lingers

An oft-touted benefit of New York State's open testing law is that it places everyone on an equal footing with those who now buy test coaching or specialized instruction from private firms. Coaching is a \$10 million industry, with more than 50,000 students annually seeking advantage over their untutored peers.

The issue is mired in controversy because the test authors—such as the Educational Testing Service (ETS) and its corporate brethren—have long maintained that coaching will not improve student scores. The exams tax accumulated wisdom, they say. Thus, no inequity exists because some can buy the coaching (typically \$200) and others cannot.

The dispute was supposed to have been resolved with a 1976 study of the coaching industry by the Federal Trade Commission (FTC) in Washington. Seduced by industry assurances that coaching was worthless, the FTC investigated the advertising claims of two coaching firms; one claimed it could increase scores on the standard college entrance exam by 16 percent. To its sharp surprise, the FTC's Boston regional office discovered the scores of some students did increase after they had been coached. So, in September 1978, the Boston staff sent its report to Washington with the recommendation that test authors, and not the coaching industry, be investigated.

ETS was hardly pleased by these results, so officials there offered some coaching of their own to the FTC Washington office. "We made contact with the people in Washington about our understanding of the way the report came out," says Robert Solomon, the vice president of ETS. "We expressed reservations about its release."

Harassed by the company and nervous about a document lending credence to a suspect industry, the FTC Washington office sat on the report for 8 months. Finally, the Boston attorney who directed the study, Arthur Levine, resigned from the agency and sued it under the Freedom of Information Act to compel the report's release. When the FTC ultimately released an abbreviated copy, it attached a cover letter pointing to "several major flaws in the data analysis, making the results unreliable." A reanalysis in Washington, however, did find that coaching "can be effective for those who do not score well on standardized tests."

Even stronger challenge to the position of the test authors is contained in an internal ETS report prepared in 1972. Lewis Pike, who was then an ETS official and who is now with the National Institute of Education, conducted an experiment in short-term instruction for the Scholastic Aptitude Test. Coached students showed score increases of as much as 13 percent (5 percent is the standard error of estimate). "Our studies showed a concentrated review of math and repeated practice with sample questions substantially raised student scores," says Pike. "Performance was particularly improved on the more complicated test items, which ETS uses more frequently now."

Pike's results were disturbing to his superiors at ETS, he says. Solomon responds that the instruction provided by Pike was simply more intense than coaching schools typically provide.—R. JEFFREY SMITH