

FairTest:

The National Center for Fair & Open Testing

342 Broadway Cambridge, MA 02139 (617)864-4810 (617) 497-2224 (fax)

Website: www.FairTest.org, e-mail: fairtest@fairtest.org

[Back to FairTest's Home Page](#)

[Back to FairTest's Main University Page](#)

Computerized Testing: More Questions Than Answers

Despite many unresolved technical and equity-related problems, test-makers are plunging headlong into new computerized methods of administering multiple-choice exams. Unfortunately, simply automating bad tests does nothing to solve their long-standing problems and may actually compound them.

The Educational Testing Service (ETS), for example, introduced its computerized Graduate Record Exam in October 1992, and has begun working on a computerized SAT. The National Council of Architectural Registration Boards already uses an ETS manufactured computerized test as part of its professional licensing process. Computerized tests are also part of ETS new generation of teacher tests, the Praxis Series, and the National Council Licensure Examination for nurses is only available on computer. In addition, many colleges, universities, and public school districts are using computerized placement tests.

Proponents of computerized testing praise its ability to provide immediate score reporting and faster and more flexible test scheduling (a test can be offered several times a week rather than a few times per year). They also point to the efficiency of Computerized Adaptive Testing (CAT), which provides a final score with fewer items in a shorter period of time than conventional paper-and-pencil, multiple-choice tests. In the future, computerized testing may also offer a greater scope of types of items through the use of graphics and video. But the new tests are being rushed into operation before adequate evidence of either their comparability to current exams or their fairness have been collected.

Unresolved Problems

Test-makers claims that the scores of computerized and pencil-and-paper tests are equivalent are inadequately supported. In fact, research studies find there usually is a difference. Most studies show higher scores for paper-and-pencil exams, but a few have found advantages for those who take computerized tests. These averages may mask

individual variations. Some respondents may still get lower scores even if the average score increases. Also, some types of questions perform differently on the two types of tests.¹

Computerized tests constrain test-takers compared to paper-and-pencil tests. With computerized versions, test-takers are unable to underline text, scratch out eliminated choices and work out math problems -- all commonly-used strategies. Studies also suggest that computer screens take longer to read than printed materials, and that it is more difficult to detect errors on computer screens.¹

Most computerized tests show only one item on the screen at a time, preventing test-takers from easily checking previous items and the pattern of their responses, two other practices known to help test-takers. Scrolling through multiple screens does not allow side-by-side comparisons.

Test-takers with the ability to manipulate computer keys rapidly may be favored by long passages that require reading through many screens.

Test-makers may try to use computerized exams to circumvent Truth-in-Testing disclosure requirements. ETS has not revealed how it intends to continue making test questions and answers available to university admissions test-takers.

Computers may worsen test bias. The performance gap which already exists on multiple-choice tests between men and women, ethnic groups, and persons from different socioeconomic backgrounds could widen as a result of computerized testing.

Schools with large minority or low-income populations are far less likely to have computers, and poor and minority children are much less likely to have computers at home^{2,3}. White students are more likely to have earned computer science credit than either African American or Hispanic students³.

The additional cost of computerized tests is certain to have a large effect on who chooses to take them. Poorer students are unlikely to take the computerized GRE, for example, because it costs nearly twice as much as the paper-and-pencil version.

Girls may be adversely affected by computerized tests. A much greater number of females than males report no school access to computers, no computer learning experiences, and limited knowledge about computers³. In addition, computer anxiety is much more prevalent among females than males⁴, with Black females reporting the greatest anxiety⁵.

SOURCES:

¹ Bugbee, A.C. & Bernt, F.M. Testing By Computer: Findings in Six Years of Use 1982-1988, *Journal of Research on Computing in Education* (Vol. 23, #1, pp. 87-100, 1990).

2 Sutton, R. Equity and Computers in the Schools: A Decade of Research, *Review of Educational Research* (Vol. 61, #4, pp. 475-503, 1991).

3 Urban, C.M. Inequities in Computer Education Due to Gender, Race, and Socioeconomic Status, (Exit Project, Indiana University, 1986).

4 Moe, K.C. & Johnson, M.F. Participants Reactions to Computerized Testing, *Journal of Educational Computing Research* (Vol. 4, #1, pp. 79-86, 1988).

5 Legg, S.M. & Buhr, D.C. Computerized Adaptive Testing with Different Groups, *Educational Measurement: Issues and Practice* (Summer, 1992, pp. 23-7).

FairTest Examiner

ETS and Test Cheating

Winter 1997

Despite much publicized efforts by the Educational Testing Service (ETS) to discourage cheating on its exams, including the persecution of apparently innocent students (see *Examiner*, Summer 1994, Summer 1992, and Spring 1992), the testmaker ignored internal warnings about security problems in its computer adaptive Graduate Record Exam (GRE), according to documents recently made public in a U.S. district court. Unlike pencil-and-paper tests, in which large numbers of students are administered identical exams simultaneously, computer adaptive tests are individualized through the automated selection of questions from a pool of items based on how each student answered the previous ones.

To protect its image and profits in the rush to introduce the new type of test into the marketplace, ETS also misled the New York State legislature about how the system actually worked.

These revelations are included in depositions submitted by top ETS officials in support of their company's lawsuit charging the Stanley H. Kaplan Educational Center with copyright violation for compiling a list of items on the computer adaptive GRE (see *Examiner*, Winter 1994-95). Kaplan executives never made the questions public, but simply announced they had them and met with ETS to demonstrate how easy it was to obtain the items.

According to court documents, a year before the November 1993 introduction of the computer adaptive GRE, two ETS test developers wrote about their fear [that examinees] will remember questions

and reveal them to their friends or to a coaching school, and that a group of examinees [might] memoriz[e] subsets of the pool and combin[e] their knowledge. Another test designer noted, Further research clearly remains to be done in this area if adaptive testing is to become a secure alternative to conventional paper-and-pencil testing, and that [N]o amount of exposure control can eliminate the possibility of an organized group memorizing the items and sharing them with others.

Disregarding those strong concerns, ETS administered a single pool of GRE questions from November 1993 through September 1994 and a second pool for the next three months, until the Kaplan disclosures. During this period, in response to questions at a New York legislative hearing where FairTest and others testified about potential problems in computerized exams, ETS reassured the State Senate, The GRE program has chosen to use many pools of questions simultaneously. That statement was not true, at least for many more months.

In fact, in her deposition, ETS President Nancy Cole admits it is now clear, in retrospect, that back in November of 1994, one or two good memorizers could have given a substantial advantage to a subsequent test-taker. Dr. Cole notes two major concerns: First, that a relatively small number of people were able to come up with so many questions; second, that there was such a high degree of overlap among the tests that were taken by the Kaplan test-takers.

After Kaplan showed ETS officials portions of 200 questions test-takers had memorized, the testmaker first suspended the computer adaptive GRE and then limited its administration for several months.

Today, ETS claims to be using a sufficiently large number of item pools to ensure that a team of memorizers cannot compromise test security, but this claim has not been subject to outside validation. Given ETS history of misleading statements on the topic and the strong incentives for test-taker collaboration, the possibility of items becoming

available to select few remains very real. Kaplan defends its actions as promoting the public interest, relying on ETS admission that the incident led to improved test design practices.

Indeed, the desire to guarantee all test-takers equal access to previously administered questions is one reason New York recently extended some truth-in-testing provisions to computer adaptive exams (see [Examiner, Fall 1996](#)). Concerns raised by the Kaplan-ETS dispute also help explain why the testmaker was unable to meet its goal of phasing out the less profitable GRE written test this year. As the controversy about many facets of computerized testing continues, the pencil-and-paper GRE, with a fresh set of items used at each administration, is likely to remain available for the next few years.

See also our [fact sheet on computerized testing](#).

August 8, 2002, Thursday

NATIONAL DESK

Officials Link Foreign Web Sites to Cheating on Graduate Admission Exams

By JACQUES STEINBERG (NYT) 872 words

An undetermined number of students in China, Taiwan and South Korea were able to raise their scores substantially last year on the verbal part of the most widely used entrance exam to American graduate schools by logging on to Web sites in those countries that post questions and answers memorized by previous test takers, test administrators said yesterday.

After uncovering the Chinese- and Korean-language Web sites and assessing their effect on scores, test administrators suspended the electronic version of the tests, known as the Graduate Record Examinations, which have been taken by 55,000 students annually in those countries since the late 1990's. Because the tests were given at testing centers six days a week, the questions were regularly reused, making the tests susceptible to such cheating.

Now, the test will be given in those three countries only two days a year, in November and March, and on paper, to guard the security of the questions, which will be used only once. Testing officials said they had not found a similar problem in any other country, including the United States.

An investigation, by the Educational Testing Service, which designs the exam, was prompted in part by the concerns of some American college deans that the high verbal scores of some Asian students did not match their English fluency.

Worldwide, nearly a half-million students a year take the exams, which measure verbal and math skills. Testing officials said they did not know how many of the students in China, South Korea and Taiwan had cheated but that it would have taken a substantial number to cause the average scores to increase as much as they did.

In China, for example, the average score on the verbal section of the exam rose 100 points in a period examined last year, on a scale of 200 to 800, testing officials said. In South Korea and Taiwan, the scores increased by 50 points.

Though the tests given in those countries often draw on the same pool of questions as the exams in given in the United States, which are also administered electronically nearly every day, testing officials said that they had not uncovered similar Web sites or unusual gains anywhere else. Testing officials said that the Web sites were probably the work of students, not bootleg businesses.

Still, Carole Beere, the chairwoman of the Graduate Record Examinations Board, which oversees the tests, said, "We are now monitoring the Web very aggressively."

Tom Ewing, a spokesman for the Educational Testing Service, said the company had rarely come upon evidence of such sophisticated, coordinated efforts to disseminate its questions and answers. In 1996, a California man was charged with recruiting people on the East Coast to take standardized tests, including the Graduate Management Admission Test, which most business schools require, only to relay the answers to students on the West Coast, in part by using pencils on which the answers were written in code.

More recently, Mr. Ewing said, the service discovered that a question on a high school Advanced Placement exam was being circulated among students via e-mail.

Though Dr. Beere said she remained confident in the validity of the exam in the United States and elsewhere, some critics argued that the breach uncovered in Asia was but one sign of the perils of administering standardized tests in the age of the Internet -- and of putting too much stock in the results of any test.

"It's not called the World Wide Web for no reason," said Robert A. Schaeffer, the public education director for the National Center for Fair and Open Testing, an advocacy group. "Those sites, even if they exist in another language, are accessible to students in the U.S. as well as in China. And there are many students in the U.S. who speak Chinese."

Dr. Beere said that this year her board had ordered a far-ranging investigation of the validity of test results in 40 countries, including the United States, and had found evidence of cheating in only those three.

Testing officials said the computerized exams drew on a rotating series of questions, often different for students taking the exam next to each other. After a certain period -- which the testing service refused to disclose, but is believed to be at least several weeks -- a question is removed from the pool permanently.

Testing officials said they were able to confirm the cheating by noting that the scores on certain questions had risen significantly the longer the question was in circulation.

Dr. Beere reported the results of the investigation to deans at more than 1,000 graduate schools in a letter mailed late last week. Its contents were first described yesterday in an online edition of *The Chronicle of Higher Education*.

In the letter, Dr. Beere, who is also the associate provost for graduate studies and outreach at Northern Kentucky University, said the board would work to "restore the confidence and trust that you have in the scores from applicants in the affected regions."

She also reminded her colleagues to view the scores of any applicant in the context of the entire application.

Copyright Registration for Secure Tests

Format Note

Table of Contents

- [What Is A Secure Test?](#)
- [Registration Procedure](#)
- [What To Bring](#)
- [Deposit Requirement](#)
- [Print Format](#)
- [Machine-readable Format](#)
- [For Further Information](#)

WHAT IS A SECURE TEST?

For the purpose of copyright registration, a secure test is a non-marketed test administered under supervision at specified centers on specific dates, all copies of which are accounted for and either destroyed or returned to restricted locked storage following each administration. A test is "nonmarketed" if copies are not sold and the test is distributed and used in such a manner that ownership and control of copies remain with the test sponsor or publisher (37 C.F.R. 202.20 (b) (4)).

REGISTRATION PROCEDURE

To enable copyright claimants to preserve the confidential nature of "secure tests," the Copyright Office provides in its regulations alternative procedures for registering secure tests and for depositing copies of these tests. To make an appointment with an examiner for the alternative procedure, **telephone the Literary Examining Section at (202) 707- 8250** and set up a mutually convenient time to have the material examined.

Once the appointment is made, hand carry the secure test material to the Copyright Office, Public Information Office, Room 401, Library of Congress, James Madison Memorial Building, 101 Independence Avenue, S.E., Washington, D.C. An information specialist will notify the Literary Section when you arrive and an examiner will meet with you. The examiner will make a preliminary examination of the complete test in your presence and immediately return the test to you, retaining adequate identifying portions of the deposit of the work to create an archival record. (See Deposit Requirement below.) Later, after the claim is fully examined and any problems resolved, a certificate of registration will be mailed to you.

WHAT TO BRING

Copyright registration of a secure test requires three elements:

1. A completed Form TX application
2. A nonrefundable filing fee of \$30* for each registration plus a secure test processing charge of \$60 per hour (checks should be made payable to Register of Copyrights)
3. A deposit of the work for which registration is sought (see Deposit Requirement below)

***NOTE:** Copyright Office fees are subject to change. For current fees, please check the Copyright Office website at www.copyright.gov, write

DEPOSIT REQUIREMENT

Print Format

For secure tests and answer material in print format, the deposit consists of:

- one complete copy (to be returned to the applicant), and
- one set of identifying material (to be retained by the Copyright Office)

The identifying material deposited with each test must constitute a sufficient archival record of the actual test; its sufficiency is determined by the Examining Division. Generally, the identifying material should consist of a photocopy of the title page of the test booklet and a photocopy of the last page of questions in each booklet with all except a narrow diagonal strip blanked out. The amount of the visible portion must be great enough to reveal that the page contains copyrightable matter. The page number must be visible on the photocopied sheet. When slides accompany the test material or when the test consists entirely of slides, the deposit in addition to the above requirements for textual matter, if applicable, consists of:

- one complete copy of the slides (to be returned to the applicant), and
- one set of identifying material (to be retained by the Copyright Office) that includes:
 - a brief written description of what is depicted in the slides and the total number of slides in the work; and
 - either a single slide from the set showing a copyrightable illustration (not text); or a photograph of a slide showing a copyrightable illustration; or a slide or photograph of a slide with all except a narrow diagonal strip blocked out. This should resemble the slit-masked copy deposited for text. It must be obvious from the portion of the slide or photograph deposited that the slide contains copyrightable illustrative material.

Machine-readable Format

For secure tests that are administered in machine-readable format or for secure tests that are administered using a traditional print-copy booklet but whose contents are taken from a larger, automated database, deposit for examination purposes consists of:

- one complete copy of the test or the database (to be returned to the applicant)

or

- 50 unmasked, complete pages of the questions or contents of the test or database (may be returned to the applicant)

These 50 pages may be either the first 25 and last 25 pages of the test or database of test questions; or they may be 50 pages from anywhere within the test or database of questions. (See Notes 1 and 2.)

NOTE 1: The applicant should decide whether he or she prefers a permanent identifying-materials deposit that includes the lesser quantity of test contents normally applied to traditional, print-format tests; or, whether he or she prefers the greater quantity of test contents applied to the category of automated databases.

NOTE 2: The applicant may also include within the registration any identifying material representing computer program authorship that may be owned by the applicant and may be considered part of the secure test work. It is not necessary, however, to register the computer program associated with an automated secure test at the same time the secure test is registered.

For permanent retention with the Copyright Office, the deposit consists of:

- a photocopy of the title page plus a photocopy of the masked or blocked-out last page of the questions showing at least some copyrightable contents, the page number, if any, and the number of the last question;

or

- a photocopy of 50 masked or blocked-out pages of the questions or contents of the test. This must be a masked deposit of the same 50 pages submitted for the actual examination.

Unusual test formats may require different identifying material. In such cases, call or write the Copyright Office to discuss particular needs or circumstances.

Warning: The Copyright Office cannot guarantee the security or confidentiality of test materials mailed to the Office. Tests received in the

FOR FURTHER INFORMATION

To speak to an information specialist, call (202) 707-3000 (TTY: (202) 707-6737), Monday through Friday, 8:30 a.m. to 5:00 p.m., eastern time, except federal holidays. Order forms and other publications from:

Library of Congress
Copyright Office
Publications Section, LM-455
101 Independence Avenue, S.E.
Washington, D.C. 20559-6000

or call the Forms and Publications Hotline 24 hours a day at (202) 707-9100. Most circulars (but not forms) are available via fax. Call (202) 707-2600 from a touchtone phone and follow prompts. Access and download circulars, forms, and other information from the Internet at : www.copyright.gov

Rev: July 2002

Format Note:

This electronic version has been altered slightly from the original printed text for presentation on the World Wide Web. For a copy of the original circular, consult the [pdf version](#) or write to Copyright Office, 101 Independence Avenue S.E., Washington, D.C. 20559-6000.

[The Library of Congress](#)

17-May-2002