

## GUEST COLUMN

Periodically we invite guest columnists to present topics of interest to education stakeholders. The following article was written and contributed by Dr. Samuel E. Krug, CEO of MetriTech, Inc.

### Testing Concerns in the 21<sup>st</sup> Century

In one form or another, testing is quite ancient. As a systematic endeavor, however, psychometrics extends back only to the late 19<sup>th</sup> century, emerging about the same time that psychology began to differentiate itself from philosophy and emerge as an empirical science in its own right. As it evolved, testing ultimately came to represent a “core technology” of psychology.

The essential notion of a test is that of a standardized behavior sample administered and scored so as to produce results that are consistent and accurate. When well-developed instruments are used correctly, they produce information that reduces uncertainty in important decisions made about people. On the other hand, when well-developed tests are used incorrectly or poorly constructed tests are used at all, the information is less than informative and may even be destructive. For more than half a century at least, professional associations have been concerned about test use and misuse (Committee on Ethical Standards for Psychologists, 1950). Their interest stems not from academic curiosity but from the realization that there is significant potential for harm when tests are misused.

Testing, which was until only a few decades ago a cottage industry, has emerged at the beginning of the 21<sup>st</sup> century as a multi-billion dollar enterprise. Test products are used widely in schools, clinics, and the workplace. They are used to determine whether students have mastered the required content and, at the same time, to evaluate whether schools have taught the required content. Clinically oriented tests are used to determine the severity of symptoms, the optimum approach to managing them, and also the progress of therapy. Often they provide the documentation for third-party reimbursement. Organizations use tests to select new employees who have the requisite knowledge, skills, abilities, and interests for the job and to identify previously unexplored talents among existing employees. The military makes extensive use of tests for placement.

### Special Challenges of Testing in the Computer Age

Perhaps no external technologies have had as much impact on testing as computers and the Internet. Although adaptive testing is at least as old as Binet’s work, the availability of inexpensive hardware and the connectivity afforded by the world-wide web made adaptive testing a more economical and large-scale activity than had been previously possible.

Substitute “misused” for “misunderstood,” and H. L. Mencken was exactly correct when he said: “Any man who afflicts the human race with ideas must be prepared to see them misunderstood.” This has certainly been the case with the idea of computer-based testing. With paper-and-pencil testing in high stakes situations, perhaps the program sponsors’ greatest security concerns involved a wayward test book or

examinees seated too close together, making answer sharing possible. Now, instead of stealing a few hundred items, hackers can grab the entire electronic item bank if security is breached. This has happened; it is not a theoretical possibility.

Internet “companies” offer professional test takers the opportunity to “receive your certification at home without spending your time and effort!” (i.e., actually bothering to take the test). Of course, some money is involved. While this might be viewed as an efficient approach from the test taker’s point of view, it would seem to violate the “standardized” requirement essential to the definition of testing in a fundamental way.

Then, too, the electronic connectivity and inexpensive storage available to the computer-based testing program is available also to test takers in the form of cell phones, programmable hand-held devices, and other instruments that make old-fashioned crib sheets and copying answers from the person at the next desk outdated. These devices also help organized groups of test takers cheat in novel but no less destructive ways to assist others taking the test later.

Of course, it is not only in computer-based testing that the concern about test misuse exists. The increased consequences for schools and staff of the educational testing requirements of No Child Left Behind raises concern as well. The pressure to meet annual improvement targets has led to documented cases of school staff changing answers to test questions, for example. Less obvious, but still important, are issues such as reassigning students to different buildings to keep numbers in reportable groups below minimum numbers. A recent study of state testing directors reports that following up on stolen test booklets or checking out allegations of improper coaching or collusion occupy nearly twice as much of their time as they did in 2001 (Sorensen, 2006). Texas recently assigned 15 people to examine allegations of test misuse, triple the number previously involved in this activity (Jacobson, 2006).

## Conclusion

Tests and psychometrics contribute importantly to decisions we make about people. The need to ensure that they remain standardized and that the information they provide is reliable and valid is no less important today than when tests first became scientific instruments. Misunderstandings regarding their purpose, problems in their administration, or confusion regarding the meaning of test scores limits their usefulness significantly. All stakeholders need to be aware of these problems and take appropriate measures for remedying them.

This article appeared in its original form in *The Score*, January 2007, vol. 29 (1), 1-3, Washington, DC: American Psychological Association.

## References

- Committee on Ethical Standards for Psychology. (1950). Ethical standards for the distribution of psychological tests and diagnostic aids: Code of standards for test distribution. *American Psychologist*, 5, 620-626.
- Jacobson, L. (2006). Probing test irregularities. *Education Week*, 26(1), 28.
- Sorensen, D. (2006). *2006 state education test security survey results*. Midvale, UT: Caveon Test Security.

