

Cheating on Large Scale Assessments: A Publisher's Perspective

Introduction

Consider the following quotations:

From the *Cleveland Plain Dealer*: "According to interviews with a dozen local teens, some test aids were taped onto the legs of school girls, just above their shirts' hemlines. Others were placed in folders with see-through covers, then plopped on desks in plain view. Many cheat sheets were attached to pens or the insides of water bottles or slipped into calculator covers. All are techniques perfected during tests taken throughout the year."

Or this, from the *Los Angeles Times*: "The State Department of Education has determined that adults apparently altered students' standardized test results in two classes at a Pasadena elementary school to improve scores... A high number of answers in two, third-grade classes were changed from wrong to right, an analysis determined. The findings will invalidate the school's gains."

A final example: "...tomorrow is the first day of the examination...it consists of a long passage which the candidate has not seen. This passage is printed on the examination paper, and it would naturally be an immense advantage if the candidate could prepare for it in advance. For this reason, great care is taken to keep the paper secret."

The report continues by describing how a student gained access to the room where the examination paper was kept, and copied it, the day before its formal administration.

Unlike the previous two quotations, in this instance we are not dealing with a state assessment program, but rather a competition for the prestigious Fortescue Scholarship. It was not a multiple-choice or a constructed response exam, but a performance assessment. The students were required to expound on a large passage of Greek translation, in this case from the historian Thucydides.

Unlike the two previous examples, this was not published earlier this year, but in June 1904 in London's *Strand Magazine*. The source was "The Adventure of the Three Students," a story by Sir Arthur Conan Doyle. Fortunately, in this instance, Sherlock Holmes happens to be in Cambridge at the time and solves the mystery. Unfortunately for us, the great detective has retired and is currently unable to help State Departments of Education and local education agencies deal with similar problems.

I use this example to make the simple point that cheating on high stakes tests is certainly not a new phenomenon.

The focus of this presentation will provide a publisher's perspective on cheating on large-scale assessments. In addition to providing a general perspective, I will also comment on the prevalence of the problem, what I believe to be the real issue, and what can be done to reduce the problem.

A disclaimer before I begin: I am providing you with **one** publisher's perspective, and though I suspect that many other publishers may come to similar conclusions, it is important to point out that I alone am responsible for my comments.

Perspective

The topic of this symposium is “Cheating on Large-Scale Assessments.” Large-scale assessments virtually always rely upon some type of standardized test. Not necessarily a norm referenced test, because the majority of state large-scale assessments are customized, though they may include a norm-referenced portion.

What is a standardized test? Simply a test developed, administered, and scored under standard conditions. Hence, we do not want to **d**estandardize any of the conditions related to the way we administer or score the test. If inappropriate activities occur, the test may become **d**estandardized and the information provided by it **d**evalued.

When can such activities occur?

1. Prior to the test being administered (as part of test preparation).
2. During the administration of the test.
3. After the administration of the test, both prior to and during the scoring process.

Allow me to provide some examples:

1. There have been numerous articles written concerning test preparation. Among the most useful are Mehrens’ and Kaminski’s 1989 piece in *Educational Measurement: Issues and Practices*. Many of you may be familiar with it and may be equally familiar with the seven-point continuum which runs from the most ethical (giving general instruction on direct objectives without referring to

the objectives that the standardized test measure), to the most **unethical** (providing practice or instruction on the test itself).

2. With respect to activities relating to the actual administration of a standardized test, virtually all test publishers provide very explicit guidelines which accompany the test materials. Similarly, customized state assessments are virtually always accompanied by even more specifically focused guidelines, and in many instances the tests are administered under semi-secure or secure conditions. Also included in both publishers' and states' guidelines are instructions relating to test accommodations for disabled and LEP students.
3. With respect to what happens after the administration of the test (both prior to and during the scoring process), it is important to note that there may be considerable variability between what a local school district using an off-the-shelf standardized achievement test does, and how that same district handles a mandated statewide assessment. The reason for these differences is that the nationally normed standardized test may be scored by the test publisher, by the local district itself, or sent to a third party for scoring. This opens numerous possibilities for "non-standard" activities.

With virtually all large-scale state assessments, however, all tests are scored uniformly, and very explicit instructions, including those relating to collecting, checking, packaging, and transporting the tests and/or answer sheets, are usually included.

So with respect to this publisher's perspective, the easiest way to deal with the situation is fairly simple:

1. If you think an activity may be inappropriate, don't do it.
2. Make sure you carefully read the guidelines before doing anything.
3. Carefully follow the guidelines.

How Prevalent is the Problem

I've already provided you with a number of recent and not so recent quotations. Consider one more if you will:

- ◆ From the *Sacramento Bee*: "It was the hottest series of numbers to hit California since the \$104 million super lotto. Up and down the state this week, people clamored to see California's first academic ranking of its nearly 7,000 schools... But in the new high stakes educational climate where scholarships for kids, bonuses for teachers, and job security for principals are all on the line—there is a potential for something completely unintended: cheating. And it's not just for kids anymore. It happened in New York, it happened in Texas, and now it's happened in Southern California, where thirteen teachers were found this month to have shared exam questions with their students before last spring's standardized testing."

We have all seen the blazing headlines and the news articles about cheating. Bear in mind, however, that newspapers don't play up honesty. What you see or hear reported is not necessarily indicative of what is happening throughout the United States. What it does represent is those instances where cheating has occurred. What you don't hear reported are the 99 percent of test administrations that go as intended.

In some instances, the fear of cheating has reached such historic proportions that it's beginning to sound like the McCarthy era all over again. Consider what recently occurred in New York City. A year after the City's special schools investigator issued a finding that dozens of teachers and two principals in thirty-two schools had helped students cheat on standardized tests, a former school inspector who reinvestigated the allegations determined that twenty-three of those teachers had, in fact, done nothing inappropriate.

Another personal example might also prove useful. As some of you know, I had ultimate responsibility for the School District of Philadelphia's assessment program through much of the 1970s. During those days we annually tested virtually every student in the School District. That's in kindergarten through grade twelve—approximately a quarter of a million students per year. During my eight and a half years heading the Office of Research and Evaluation, we discovered only two instances of deliberate attempts to cheat. One involved a principal who endeavored to put pressure on some of his teachers, and who was turned in by the teachers in that school. The other instance was a single teacher (in another school several years

later) who provided the answers to her students. In this case, when it was discovered by the other teachers, they reported her to the central administration.

What is my conclusion concerning the prevalence of the problem? Given the tremendous increase of large-scale assessment, the problem is indeed more prevalent than a decade ago, but not nearly as common as the media would have us believe.

The Real Issue

Psychologists tell us that there are some individuals who are just naturally dishonest. The most publicized kleptomaniacs have been extremely wealthy individuals. These findings aside, however, it is extremely rare that individuals risk cheating on a test that doesn't have consequences. At the risk of having everyone in the audience say "duh," I'll state the obvious. "It's not the scale, it's the stakes." Large-scale testing is not the culprit; it is large-scale testing with high stakes attached to the results, that, in some instances, has resulted in cheating.

I would further suggest to you that in these instances what has occurred is that the test results are being viewed as an end unto themselves rather than as information about achievement.

What Can Be Done To Avoid The Problem

Imagine, if you will, a solution spectrum. At one end, there are a number of solutions that might fall into the "get tough" category. These include (1) better

security, (2) using technology to catch the perpetrators, and (3) stiffer penalties for those who are caught. Will these help? Probably, to some extent.

Of course, at the other end of the testing continuum there are groups such as “No Test,” which suggest fixing the problem by doing away with tests altogether. Since this is quite unlikely in the current political, social, and educational climate, we might wish to consider a middle ground. This middle ground calls for placing the role of assessment in context.

For example, Joe Nathan, Director the University of Minnesota’s Center for School Change based at the Humphrey Institute of Public Affairs, has found that there are nine key elements that schools, schools districts, and states can employ in designing accountability and assessment programs. Nathan indicates that six of these elements are **vital** in assuring appropriate approaches. They are:

- ◆ Clear and measurable goals for each school.
- ◆ Goals that are understood and supported by families, students, and educators.
- ◆ A variety of measures, including standardized tests and performance measures.
- ◆ The measurement of all students’ work, including that of special education students.
- ◆ Assessments that measure the progress of students who don’t speak English at home.
- ◆ The use of assessment information to shape school improvement efforts.

Among the other three elements, which Nathan deems “valuable” (but not vital), is one which I should suggest would be in the earlier category. And that is: “Creating a parent/educator/community committee to supervise assessment efforts.” I personally think that that such a group would be invaluable in helping forestall any irregularities in local assessment programs.

I would also like to suggest that we pay more attention to a series of recommendations made by the National Research Council report on assessment, and to the most recent revision of the APA/AERA/NCME *Standards for Educational and Psychological Testing*. In particular, states or local districts that make use of high stakes assessments should be sure that:

1. Students receive adequate notice of the test and its consequences.
2. Students have an opportunity to learn the knowledge and skills being tested, meaning that the test must be aligned with the curriculum.
3. The process for setting passing scores should be documented and evaluated.
4. Students should have equal access to any specific preparation for taking the test.
5. Students who risk failing the test should be advised in advance and provided with appropriate instructional remedial help that will improve their chances of passing it.
6. Students should have multiple opportunities to retake the test.

Finally, let me suggest an area that receives a great deal of lip service and little else. And that area is training, especially training teachers, principals, education policy makers, and the media to understand the appropriate role of assessment in education.

We could probably spend an entire session talking about the paucity of assessment training and the resulting misuse of assessment information. But that's a topic for yet another time, however.

Conclusion

In conclusion, let me state that CTB/McGraw-Hill, and I suspect most other test publishers, publish a list of explicit practices and procedures as part of their test administration guides. Such procedures reflect general ethical principles and standards, as typically reflected in a variety of well-known documents and position statements published by such organizations as AERA, APA, and NCME. Procedures reflect such areas as test security, pre-test activities (which include test preparation activities), testing conditions (including specific testing procedures, testing in the classroom, testing materials, directions, and monitoring), and post-test activities.

Test publishers and their partners—the state departments of education and local school districts with which they work—have collaborated closely in the past, and continue to do so today in the formulation of explicit procedures to guide the

administration of all assessments, including large-scale assessments, especially those with high-stakes consequences.