



8-1-1976

Accountability and Test Security

A. J. Polemeni

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: <https://commons.und.edu/tl-journal>



Part of the [Scholarship of Teaching and Learning Commons](#)

Recommended Citation

Polemeni, A. J. (1976) "Accountability and Test Security," *Journal of Teaching and Learning*: Vol. 2: Iss. 2, Article 2.

Available at: <https://commons.und.edu/tl-journal/vol2/iss2/2>

This Article is brought to you for free and open access by UND Scholarly Commons. It has been accepted for inclusion in *Journal of Teaching and Learning* by an authorized editor of UND Scholarly Commons. For more information, please contact und.common@library.und.edu.

Accountability and Test Security

Anthony J. Polemeni

This bicentennial year provides an excellent opportunity for modern educators to reflect upon the relative significance of educational developments over the last decade in these United States. Perhaps the most pervasive influence over this period is that of educational accountability. The far reaching impact of accountability is recognizable in virtually every philosophy, theory and practice of learning. This is true to a large degree because it is the purpose of modern education to place itself under intensive scrutiny to a degree never before dreamed of in our history.

In general, the effect of this intensification of focus on progress is beneficial. It motivates teachers to ask more questions and schools to establish evaluation programs which are more definitive. Various attitudes which have resisted change for decades are justifiably challenged, and some of the rigidity that is found in academic life is replaced by more wholesome and fresh processes of creativity, as new needs are met in the everyday affairs of our busy schools.

On the other hand, there are certain ways in which even accountability itself is less than beneficial. The anxiety that may be generated by accountability in certain situations, such as testing, should be, and is, of great concern to many educators. Over-reaction to the concept of accountability is of some importance in any large-scale testing program. Enormous pressures are generated upon school staffs who are subjected to hypercritical or conflicting demands for increased performance.

Controversy in testing is nothing new. During the past three decades, there have been various controversies such as those involving invasion of privacy, allegations of bias, and demands for culture fairness.

¹Director, Office of Educational Evaluation, New York City Board of Education.

The expansion of accountability, however, with its insatiable appetite for more relevant instrumentation and more representative sampling, is not likely to bring about an end to controversy in our lifetime. There are a great many questions to answer. For example, where exactly will we draw the line between the conflicting perspectives of norm referenced and criterion or content referenced testing? And when will we be able to say that the job in testing is done adequately for the bilingual segment of our sprawling American population? These and similar questions may make us consider whether accountability has put testing back into a stage of infancy.

Another important, although less popular, problem has arisen regarding evaluation in direct response to the pressures of accountability, and that is - test security. In the New York City Public Schools System there is a long and arduous history of dealing with certain irregularities and abuses in the administration of standardized tests. These irregularities consist of various forms of coaching children on specific test content, or even using excessive amounts of classroom time to practice test taking. Undoubtedly, the pressures mounted by parents and legislators have backed up on teachers, supervisors and school organizations to fulfill expectations that are often not well-founded or even conflicting. There are cases, for example, where the frequency of each child's classroom interaction is suggested as a criterion for teaching and learning. The quality of the interactions is usually neglected.

Everyone agrees that teacher competency must be reviewed and rated, but there are more than a dozen different schools of thought about precisely who it is that should rate whom, and what criteria and method should be used under which set of circumstances. A great deal of undesirable tension comes from this, and it can lead to coaching of students by teachers who want to survive.

Supervisors who are pressed to provide intensive rating information find that conflicting standards prevail in everyday classroom accountability. These beleaguered raters find that there is more than one public that makes demands; there are parent groups and their

standards; evaluators with agency standards; higher officials with school rankings uppermost in their thinking; and legislators who calculate per pupil expenditures. Apart from these, there are separate demands made in terms of accountability by national assessment teams. Who is accountable to whom and for what? How do these conflicting standards reflect on the testing program?

Despite conflicting opinions on the use of tests, the standardized test has yet to be replaced by other criteria for educational accountability. Without these data there are few broad frames of reference.

This leads us to our central concern in this discussion: namely, in what specific ways does the pressure of obtaining a favorable place in supervisors' ratings or school rankings influence testing unfairly, and what system should be established as a countermeasure?

Test security is violated in a number of obvious ways and in other ways which are not so obvious. Over the past few years the New York City Board of Education has regularly advised school officials against certain questionable practices concerned with the administration and uses of tests, and at the same time it has recommended approved practices. Some of these recommendations are summarized here.

Unacceptable Practices

1. Departing from directions - especially modifying time limits and prompting pupils on items
2. Employing material from standardized tests to orient pupils to test conditions
3. Familiarizing pupils with specific test content (e.g., vocabulary lists)
4. Teaching for the test by using standardized tests of any kind regardless of form, level or edition.

Acceptable Practices

1. Orienting pupils to relaxed "test-like" situations
2. Giving pupils practice in following verbal directions and working under time limits

3. Acquainting pupils with test-taking mechanics such as the use of a separate answer sheet
4. Familiarizing pupils with common item formats
5. Developing the test-taking skills of pupils such as using time properly, avoiding errors, guessing wisely and reasoning out answers.

At this point one might reasonably ask: What is the difference between acceptable and unacceptable practices as they appear above? Would not those students who were instructed in "acceptable practices" have some advantage over those who were not given such instructions? The answer is, yes.

Although it is quite clear that the difference between acceptable and unacceptable practices in testing is a matter of how specific the coaching is, in terms of actual test item content, the two approaches are not far apart.

In fact, it might be argued successfully that any pedagogical instruction which centers on standardized exam-taking skills for one rather than another group of students is unfair. The correct assertion would then be that unless all students can enjoy equal exposure to such an advantage, there should be no separate advantage to any particular group. All instruction would be set within the four walls of curriculum with testing neatly fenced out of instruction and skill development. However, we know that some test-taking instruction should be available.

Some children naturally possess more test-taking ability than others, and this often disguises true achievement differences. To reduce such differences is highly desirable. It is also true that children differ in their knowledge of the mechanics of testing, test anxiety, test sophistication and experience. If it is desirable to counter some of these differences, and we believe that it is, then we must implement programs which (a) fairly increase test sophistication to all pupils and (b) reduce irregularities and abuses in the use of the tests.

Such programs must be developed centrally, so that the widest application of similar instruction can be made a function of curriculum. This would help to reduce the anxiety of the children and thus should improve performance.

Abuses such as teaching test content and coaching are best reduced by a secure testing program which insures that:

1. Actual testing materials are in the schools only during a specific testing period and not before or after.
2. A different, but parallel, form of a test is used each year. This calls for extensive test development, but the old custom of re-using an existing form every four or five years within a large city should be avoided.
3. A wholly revised test series should be developed at least every six years. This would assure not only greater security, but it would also better provide for representativeness of the test content, current norms and curriculum relevance.

A combination of teacher training in test security and the annual development and validation of equivalent forms of standardized tests should help to lessen some of the excessive anxiety in test use and counter test abuse as a result.

Thus far in our history the standardized test remains as the central tool of accountability, and its use must be better understood and protected by all participants in the testing program. It is important for teachers to know as much as possible about what information standardized tests can and cannot provide and how the tests can help answer the questions raised by educational accountability. It is also necessary for each teacher to be directed in the proper uses and possible abuses of tests - both norm-referenced and criterion-referenced.