

# Appropriateness of Teachers' Test-Preparation Practices

W. James Popham

*University of California, Los Angeles*

*What two standards can teachers and administrators use to decide whether a particular way of preparing students for a test is appropriate? How do five commonly used test-preparation practices stack up to these standards? How do educators and school board members view these five practices?*

Faced with growing pressures to increase students' scores on achievement tests, some American educators have responded by engaging in test-preparation practices ranging from those that may be educationally debatable to those that are downright dishonest. In recent years, a number of instances have been reported regarding teachers and administrators who deliberately coached students with actual copies of a supposedly "secure" examination (Cannell, 1989). There are even reports of educators' erasing students' incorrect answers and replacing them with correct answers. Such acts, of course, stem from the increasing emphasis on student test performance as the chief indicator of an educational program's success.

## **Origins of High-Stakes Testing**

The preoccupation with student test scores as the definitive indicator of educational effectiveness was, in the main, a phenomenon of the 1980s. Although there had been attention given to the quality of student test performances prior to that decade, it was in the 1980s that attentiveness to student test performance became pervasive. In large measure, the focus on pupils' test scores stemmed from increasing incredulity on the part of American citizens that the nation's public education system was performing properly. Taxpayers, and their elected representatives, registered serious doubts that

educational tax dollars were being well spent. Spawned by such doubts, the era of educational accountability became a reality in the 1980s when state after state enacted laws requiring pupils to take annually administered achievement tests. The results of such tests were used not only to determine the quality of statewide schooling, but also (because of the widespread practice of publishing such test scores in local newspapers) to mirror the effectiveness of individual school districts and schools.

In every sense of the expression, these legislatively mandated achievement tests were *high-stakes tests* because there were significant contingencies associated with the test results. The contingencies were experienced either by the students who took the tests or by the educators who administered the tests. To students, the tests were significant because in many instances test scores were linked to high school graduation or grade-level promotion. To educators, because of the manner in which test results were publicized by the local media, high test scores were viewed as indicating an effective instructional program and, of course, vice versa. Because of these contingencies, America's teachers and administrators found themselves frequently called on to raise test scores. Candid educators will recount that the past decade's pressure to boost students' test

scores was both persistent and profound.

Discussions with American educators regarding what sorts of test-preparation practices they regard as appropriate or inappropriate lead to one inescapable conclusion: most American educators have not devoted serious thought to the appropriateness of different test-preparation practices. Given the relative recency of the prevalence of high-stakes testing in American education, it is not surprising that little attention has been given to the appropriateness of various test preparation practices. Yet, because no decrease in the pressures on educators to promote higher test scores is likely, the time has arrived for America's teachers and administrators to consider seriously what sorts of test-preparation practices are, indeed, appropriate.

It is anticipated that instructional specialists and measurement experts will sharpen the issues associated with this increasingly important issue. Yet, thus far, few writers have devoted serious attention to the topic. Among the exceptions are Mehrens and Kaminski (1989), who have supplied thought-provoking observations on the topic of test preparation.

## **Tests as Indicators of a Student's Status**

Before addressing topics related to test preparation, it is important to clarify the function of educational

---

*W. James Popham is a professor emeritus at the Graduate School of Education, University of California at Los Angeles, and is director of IOX Assessment Associates, 5420 McConnell Ave., Los Angeles, CA 90066. He specializes in educational measurement and evaluation.*

achievement tests. An educational achievement test is employed in order for us to make a reasonable inference about an examinee's status with respect to a domain of knowledge and/or skills it represents. Ideally, of course, an achievement test will sample the content domain representatively so that the level of an examinee's performance on the achievement test will serve as a reasonably accurate reflection of the examinee's status with respect to the content domain. Thus, for example, an examinee who answered correctly 80 percent of the items in an achievement test would be expected to have mastered about 80 percent of the content in the domain of knowledge and/or skills that the test is measuring. The relationship between a student's test performance and that student's mastery of the content domain represented by the achievement test, as will be seen later, is a key factor in establishing the appropriateness of test-preparation practices.

## Two Evaluative Standards

Two standards can be employed by teachers and administrators who wish to ascertain the appropriateness of given test-preparation practices. Taken together, the two standards provide guidance to educators who wish to determine the suitability of particular test-preparation activities.

*Professional Ethics:* No test-preparation practice should violate the ethical standards of the education profession. This first standard obliges educators to avoid any test-preparation practice that is unethical. Ethical behaviors, of course, are rooted not only in fundamental morality, but also in the nature of a particular profession. For example, physicians should be governed not only by general ethical principles dealing with honesty and respect for another's property, but also by a set of ethical principles that have evolved specifically for the medical profession. Similarly, educators should not engage in test-preparation practices that involve violations of general ethical canons dealing with theft, cheating, lying, and so on. In addition, however, educators must take seriously the ethical obligations that they undertake because they have

agreed to serve *in loco parentis*. Educators who serve "in place of the parent" take on an ethical responsibility to serve as *models* of ethical behavior for children.

In passing, it should be noted that when educators engage in test-preparation practices that, if brought to the public's attention, would discredit the education profession, such practices border on the unethical because, in the long term, they erode requisite public support for our schools and hence render the education profession less potent.

Thus, according to the standard of professional ethics, educators should not engage in test-preparation practices that involve such behaviors as violating state-imposed security procedures regarding the content of high-stakes tests. A growing number of states have enacted regulations so that educators who violate state test-security procedures could have their credentials revoked. Accordingly, educators should not engage in test-preparation practices that are unethical because there are potential personal repercussions (e.g., loss of credential) or professional repercussions (e.g., reduced confidence in public schooling) and, most importantly, because unethical test-preparation practices are *wrong*.

*Educational Defensibility:* No test-preparation practice should increase students' test scores without simultaneously increasing student mastery of the content domain tested. The second standard emphasizes the importance of engaging in instructional practices that are in the educational best interests of students. Teachers should not, for example, artificially increase students' scores on a test while neglecting to increase students' mastery of the domain of knowledge and/or skills that the test is supposed to reflect. An appropriate test-preparation practice raises not only students' prepreparation-to-postpreparation performance on a *test* but also raises students' mastery of *the content domain* being tested. Conversely, an inappropriate test-preparation practice raises students' prepreparation-to-postpreparation performance on the test, but not students' mastery of the content domain itself.

The result of such inappropriate test-preparation practices is that

a *deceptive* picture of students' achievement is created. The test results no longer serve as an accurate indicator of students' status with respect to a content domain. As a consequence, students who in reality have not mastered a domain of content may fail to receive appropriate instruction regarding such content. The students will be instructionally shortchanged because inappropriate test-preparation practices led to an inflated estimate of their content mastery. Such test-preparation practices, because they rob students of needed instruction, are educationally indefensible.

## Five Test-Preparation Practices

We now turn to a consideration of five common test-preparation practices. These five practices are not exhaustive in the sense that there are no other conceivable test-preparation practices. The five practices do, however, capture most of the important test-preparation options available to teachers. Some of these practices involve *special* instruction rather than *regular* classroom instruction. Special instruction consists of extra preparation sessions, during or outside class, that are devoted exclusively to the readying of students for tests. In contrast, regular classroom instruction, although its focus may be relevant to the content of a test, occurs as part of the teacher's ongoing instructional program. The five test-preparation practices are given below.

1. *Previous-form preparation* provides *special* instruction and practice based directly on students' use of a previous form of the actual test. For example, if the currently published form of a nationally standardized achievement test is being used on a statewide basis, the teacher gives students guided or independent practice with earlier, no longer published, versions of the same test.

2. *Current-form preparation* provides *special* instruction and practice based directly on students' use of the form of the test currently being employed. For example, the teacher gives students guided or independent practice with actual items copied from a currently used state-developed high school graduation test.

3. *Generalized test-taking preparation* provides *special* instruction that covers test-taking skills for dealing with a variety of achievement test formats. For example, the teacher shows students how to make calculated guesses for certain types of items or how to allocate test-taking time judiciously.

4. *Same-format preparation* provides *regular* classroom instruction dealing directly with the content covered on the test, but employs only practice items that embody the same format as items actually used on the test. For example, if the achievement test includes addition problems formatted only in vertical columns, the teacher provides practice with problems formatted solely in that manner.

5. *Varied-format preparation* provides *regular* classroom instruction dealing directly with the content covered on the test, but employs practice items that represent a variety of test item formats. For example, if the achievement test uses subtraction problems formatted only in vertical columns, the teacher provides practice with problems presented in vertical columns, horizontal rows, and story form.

### Applying the Standards

When these five test-preparation practices are scrutinized according to the standards of professional ethics and educational defensibility, only two turn out to be appropriate.

1. *Previous-form preparation* violates the educational defensibility standard because students' test scores are apt to be boosted via such special preparation sessions without concomitant rises in mastery of the content domain being tested. In addition, because in the public's view this kind of test-preparation practice may be seen as an improper instance of coaching students *merely* for test score gains, it may be viewed by some as unethical. Previous-form preparation, therefore, is *inappropriate*. (This same judgment would apply to the use of commercial test-preparation materials based chiefly on newly created "parallel" forms of a currently used test.)

2. *Current-form preparation* clearly loses out on both standards. Not only is it educationally indefensible, but such preparation consti-

tutes an outright example of cheating. Current forms of the test must be stolen or surreptitiously copied in order to be used in such potentially improper score-boosting special sessions. Educators who are caught readying their students via current-form preparation should, in fact, be caught.

3. *Generalized test-taking preparation* turns out to be an appropriate form of test preparation because such special instruction on how to take tests is, characteristically, rather brief and hence not seriously deviating from a student's ongoing education. More importantly, because such test-taking preparation readies students to cope with a number of different sorts of tests, students will be less apt to be intimidated by a previously unencountered type of test item. In a very real sense, therefore, generalized test-taking preparation sessions allow students' test performances to be more accurately reflective of their true state of knowledge and/or skill. Such preparation, if not excessively lengthy, is clearly appropriate.

4. *Same-format preparation*, although it may be ethical, is not educationally defensible. If students in their regular classroom instruction are *only* allowed to deal with the explicit item format used on a test, then those students will be far less likely to generalize what they have learned. Test scores may rise, but content mastery is less likely to rise concomitantly. Although many administrators may, because of powerful pressures to boost scores, endorse such test-preparation practices, they should be resisted because they are educationally unsound.

5. *Varied-format preparation*, in contrast, satisfies both of our evaluative standards. Students during their regular classroom instruction are given practice not only with content as it is conceptualized on the test, but also with content conceptualized in other ways. Rises in test scores will, in general, be accompanied by rises in mastery of the content domain being tested because students' generalized mastery of the content is being fostered.

In my view, then, of the five test-preparation variants, only varied-format preparation and general-

ized test-taking preparation satisfy both evaluative standards.

In an effort to discern how teachers, administrators, and school board members are apt to regard the appropriateness of various test-preparation practices, a brief survey was given to groups of workshop participants at the *beginning* of three workshops in late 1989 and early 1990. The first session was held in southern Ohio in November, 1989, and was attended by teachers and administrators from Ohio, Indiana, and Kentucky. The last two sessions were held in Los Angeles in January and February of 1990 and were attended by teachers, administrators, and school board members from southern California. It was hoped that participants would hold different views at the close of the workshops in which the topic of test-preparation practices was considered. Their views at the beginning of the workshops, however, were apt to be representative of the way that such individuals view the appropriateness of the practices as described.

All participants were given descriptions of five test-preparation practices identical to those supplied earlier in this analysis. These individuals were asked to supply anonymous judgments regarding whether each practice was appropriate or inappropriate based on brief descriptions of the two standards described earlier, that is, professional ethics and educational defensibility. In all, 172 respondents judged the appropriateness of the five test-preparation practices. Results, identified by group, are supplied in Table 1.

In considering Table 1, we see that most respondents concurred with our previous analysis that generalized test-taking preparation and varied-format preparation are appropriate. With respect to the other three preparation procedures, however, there is far from universal agreement. In particular, one is struck by the large percentages of respondents (from 6 to 36 percent) who consider it appropriate to give special instruction to students "with actual items copied from a currently used" test. Moreover, decisive majorities of all respondent groups believe that same-format preparation is appropriate even though we have con-

**Table 1****Views of the Appropriateness of Different Test-Preparation Practices**

Test-preparation practice	Percent considering appropriate					
	Midwest teachers <sup>a</sup>	Midwest administrators <sup>b</sup>	California superintendents <sup>c</sup>	California board members <sup>d</sup>	California teachers <sup>e</sup>	California principals <sup>f</sup>
Previous-Form	34	47	60	68	57	25
Current-Form	14	17	17	21	36	6
Gen. test-taking	95	96	96	97	100	100
Same-Format	52	72	83	66	64	69
Varied-Format	90	91	96	92	100	94

<sup>a</sup>n = 21; <sup>b</sup>n = 46; <sup>c</sup>n = 23; <sup>d</sup>n = 38; <sup>e</sup>n = 28; <sup>f</sup>n = 16.

cluded such preparation often instructionally shortchanges students.

It is always possible, of course, that the educators and board members who judged the appropriateness of the five test-preparation practices may regard today's uses of high-stakes tests as sufficiently noneducative so that *any* sort of test preparation should be considered appropriate. Putting it another way, not every educator would choose the two evaluative standards recommended here or, for that matter, apply them in the same way.

### Confronting the Issue

American educators are apt to find that their own test-preparation practices may not coincide perfectly with one of the five practices described. However, it is likely that the major dimensions of test-preparation practices are adequately represented by the preparation approaches considered here. More important than the five preparation practices, of course, are the two evaluative standards. Teachers and administrators must review with care their own preparation activities linked to high-stakes tests. The two evaluative standards provided here—professional ethics and educational defensibility—may prove useful in such reviews. Educators who seriously scrutinize their test-preparation activities are apt to discover that some of those activities are definitely inappropriate.

Given the scant attention that this issue has received, local and state boards of education could provide helpful guidance to educators by developing policy positions regarding what sorts of test-preparation

practices should be applauded and what sorts should be expunged.

### Note

This article was initially presented at the Forum for Dialogue between Educational Policymakers and Educational Researchers, sponsored by the UCLA Graduate School of Education and the California School Boards Association, University of California, Los Angeles, January 27, 1990. It was presented in expanded form at the Educator Awards

Conference, sponsored by the Foundations of the Milken Family, Los Angeles, May 5, 1990.

### References

- Cannell, J. J. (1989). *How public educators cheat on standardized achievement tests*. Albuquerque, NM: Friends for Education.
- Mehrens, W. A., & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent. *Educational Measurement: Issues and Practices*, 8(1), 14–22.

### State-Wide Assessment

*Continued from p. 6*

An assessment system that makes it possible to hold districts accountable for student outcomes would enable state departments of education to shift their emphasis from enforcing compliance with state rules and regulations to promoting research and service. There are indeed districts that are in desperate need of help. A sound state assessment system could help to identify such districts, and a state department staffed with people who know how to help could be part of what the state does if state officials are to achieve the goal for which they are constitutionally accountable: maintaining a thorough and efficient system of public education.

### References

- Burstein, L., Baker, E., Aschbacher, P., & Keesling, J. W. (1986). *Using state test data for national indicators of*

*education quality: A feasibility study*. Los Angeles: Center for the Study of Evaluation.

- Cooley, W. W. (1990). *Important variations among Pennsylvania school districts*. Pittsburgh: Pennsylvania Educational Policy Studies, University of Pittsburgh.
- Cooley, W. W., & Bickel, W. E. (1986). *Decision oriented educational research*. Boston: Kluwer.
- Cooley, W. W., & Bernauer, J. A. (1990). *School comparisons in state wide testing programs*. Pittsburgh: Pennsylvania Educational Policy Studies, University of Pittsburgh.
- Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18(9), 27–32.
- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. Travers (Ed.), *Critical Issues in Curriculum. Eighty-seventh Yearbook of the National Society for the Study of Education* (Part I, pp. 83–121). Chicago: University of Chicago Press.
- Resnick, L. B. (1987). *Education and learning to think*. Washington, DC: National Academy Press.