

**Test Preparation:
Considering the Appropriateness of these Activities**

A Professional Development Module for Iowa Educators

An online version of this module can be accessed at:

<http://www.iowa.gov/educate/content/blogcategory/497/920/>

Prepared for

Iowa Department of Education

Prepared by

Center for Evaluation and Assessment
The University of Iowa

© The University of Iowa, Center for Evaluation and Assessment, 2007.

This document may be copied and distributed to Iowa schools for use within the state of Iowa. Other uses are prohibited.

Table of Contents

Background	ii
Study Plan	iii
Introduction	1
Part 1: What is “test preparation?”	4
Part 2: Why do educators feel that test-preparation activities are necessary and/or important?	5
Part 3: Why should educators care about the “appropriateness” of test preparation	8
Part 4: How can the appropriateness of a test-preparation activity be determined?	10
Part 5: What should educators do when they feel pressured to use inappropriate practices or they are aware of such practices being conducted in their school?	28
Part 6: What can/should be done to increase the likelihood that scores are a true reflection of what students know and are able to do in the areas covered by the tests?	31
Closing	39
Instructional Aides:	
Guidance from Iowa Testing Programs (August 15, 2005)	A1
Scenario Activity	B1
Scenario Consensus Guide	C1
Closing Activity	D1
Closing Activity Feedback	E1

Background Related to the Development of this Module

To be effective, this module should do the following:

1. Consider the background knowledge of the intended learners and their current working environments.
2. Grasp the educator's attention right away, using examples that highlight the potential for extreme negative consequences resulting from using inappropriate activities.
3. Use a tone that is friendly, yet not condescending.
4. Use easily understood terms and vocabulary.
5. Incorporate the guidance provided by Iowa Testing Programs so as to promote a recognizably consistent message.
6. Incorporate time for interaction and the exchange of ideas and perceptions.
7. Help educators identify not only what is wrong with a particular activity, but also how the activity could be modified to make it more appropriate.

Premises regarding the intended audience for this professional development module: (Developed for use by Iowa educators)

Background

- Pre-service teacher education programs and graduate programs in educational administration provide very little guidance on sound measurement/testing practices.
- Nearly all actions taken by educators are based on "good intentions."

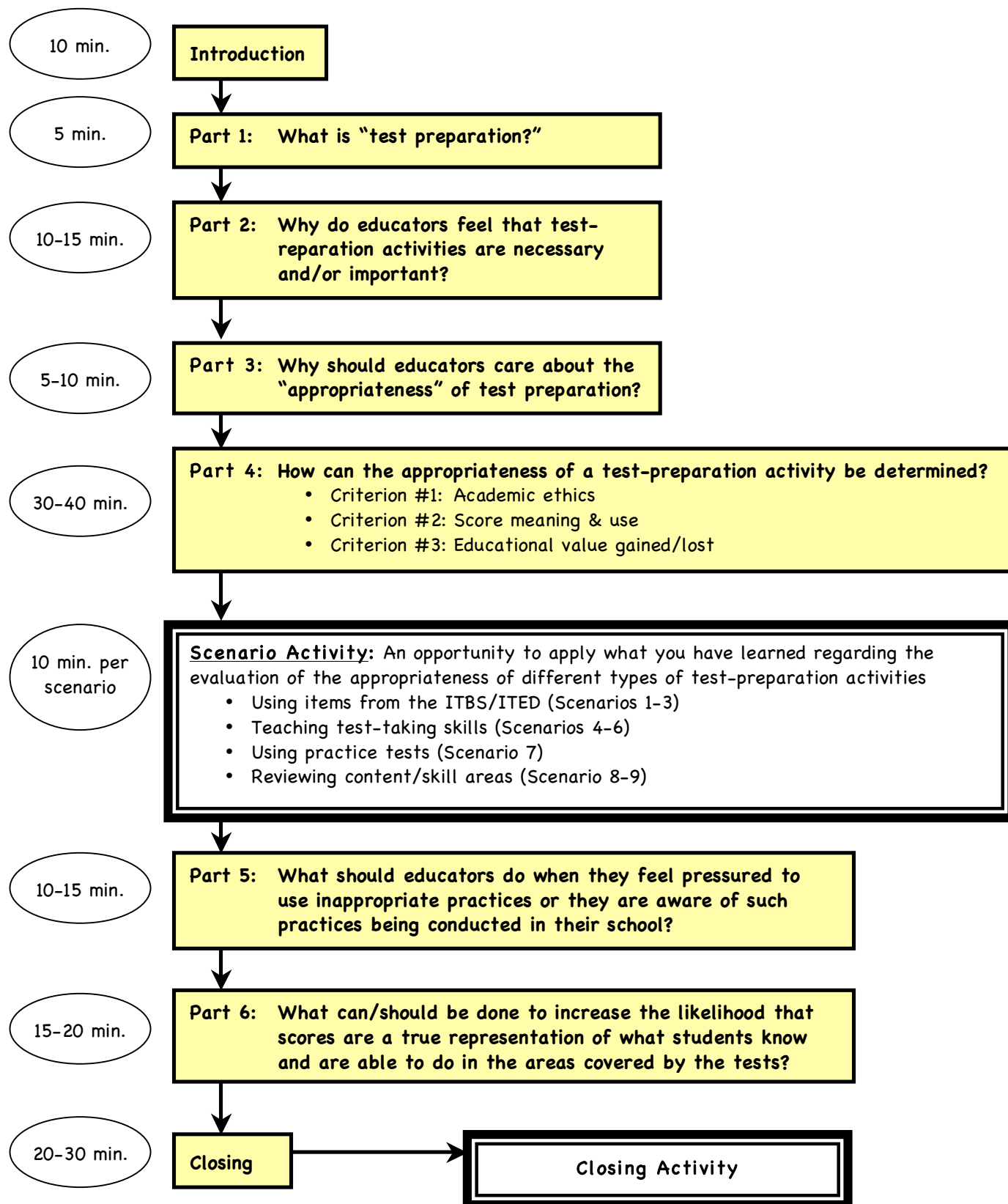
Educators as evaluators of student achievement

- Some educators believe that serving as an evaluator is in conflict with their mission of improving student learning.
- Some educators have negative perceptions regarding the *ITBS/ITED*; these perceptions are often based on personal experiences or on misinformation regarding the tests.
- Some educators believe that the manner in which a student is assessed should match completely with what the student has encountered as part of everyday instruction.
- Some educators are not familiar with critical features of the *ITBS/ITED* and/or with the purposes for which the *ITBS/ITED* is administered and how the scores can be used to inform instruction.

Testing for accountability

- Use of the *ITBS/ITED* in Iowa has historically been for low-stakes purposes, thus many educators have not previously been exposed to the pressures associated with this context.
- Some educators are inclined to respond to external pressure to increase student scores.
- Some educators do not recognize that some activities that might be appropriate for helping students to prepare for an assessment in one context (i.e., classroom test) might not be appropriate in other types of testing situations.
- Some educators have not had the opportunity to think critically about the link between increased test scores and increased student learning.

Study Plan



Total amount of time required: About 2 hours, without completing the scenarios

Introduction

Purpose

This module has been designed for use by Iowa educators to assist them in making sound decisions regarding activities associated with preparing students to take the *Iowa Tests of Basic Skills (ITBS)* or *Iowa Tests of Educational Development (ITED)*. Although these two test batteries have been used in Iowa schools for nearly 70 years, historically the scores have been used for instructional decision-making and have not been tied to high stakes for students or schools. The enactment of NCLB, however, has added school-level accountability to the ways in which the scores from these tests are being used, and it is likely that many educators have been (or will be) inclined to respond to external pressure to increase student scores by engaging students in test-preparation activities. Because some of these activities can result in inaccurate scores and produce negative consequences for students, teachers, schools, parents, and the community, it is important that educators (teachers and administrators) be able to make educationally and ethically defensible decisions about what types of activities can be used and in what contexts.

It's important to clarify that the guidance presented in the module pertains to the use of test preparation in the context of preparing students to take the *Iowa Tests*. Some activities that might be appropriate for helping students to prepare for an assessment in one context (i.e., classroom test, ACT) might not be appropriate in other types of testing situations.

Using this Module

This professional development program has been designed for use in the following types of situations.

- Independently, self-paced by individual educators
- Small-group settings, with opportunities for discussion

Throughout the module, opportunities for reflection and interaction have been incorporated. It's believed that these opportunities are essential for developing a sound understanding of the complexities associated with preparing students for testing in the context of school-level accountability.

Objectives

At the completion of this professional development module, educators should be able to:

1. Explain how criteria related to academic ethics, score meaning & use, and educational value are related to test-preparation activities.
2. Explain how the ways in which the *Iowa Tests* are used in the state have bearing on the appropriateness of a given test-preparation activity, in terms of academic ethics and/or score meaning & use.
3. Use criteria related to academic ethics, score meaning & use, and educational value to evaluate the appropriateness of a given test-preparation activity and make recommendations regarding how the activity might be modified, if necessary, to be more appropriate in light of these three criteria.

4. Identify negative consequences for students, teachers, schools, parents, and the community resulting from the use of improper test-preparation activities.
5. Conceptualize a personal philosophy regarding test preparation, which incorporates the three criteria related to academic ethics, score meaning & use, and educational value.

Components:

- Part 1 What is “test preparation?”
- Part 2 Why do educators feel that test-preparation activities are necessary and/or important?
- Part 3 Why *should* educators care about the “appropriateness” of test preparation?
- Part 4 How can the appropriateness of a test-preparation activity be determined?
- Part 5 What should educators do when they feel pressured to use inappropriate practices or they are aware of such practices being conducted in their school?
- Part 6 What can/should be done to increase the likelihood that scores are a true representation of what students know and are able to do in the areas covered by the tests?

Instructional Aids:

Print-versions of these materials can be obtained from the “Downloads” page.

Study Plan: A general outline, which summarizes the organization of this module, has been provided so that you can more easily reflect upon where you are as you progress through the module. This plan also contains estimates of the amount of time required to complete each part of the module.

Flowchart: A stepwise process for evaluating the appropriateness of any activity used for test-preparation purposes has been summarized in the form of a flowchart. This flowchart is presented in Part 4.

Guidance from Iowa Testing Programs: Iowa Testing Programs (the developers of the *ITBS* and *ITED*), in conjunction with the Iowa Department of Education, disseminated guidance on the development of district policy regarding test use, test preparation, and test security as it relates to the *Iowa Tests* (Iowa Testing Programs, August 2005). This guidance has been incorporated into this module, where appropriate, to promote a consistent message regarding these types of practices.

Scenario Activity: A set of nine scenarios accompanies this module so that you can apply what you have learned regarding the evaluation of the appropriateness of different types of test-preparation activities. These scenarios are best used after the completion of Part 4. Although you can work through these scenarios individually, it’s believed that you would greatly benefit by working in a small group because of the opportunity to interact and exchange ideas and perceptions. There is no need to work through the scenarios sequentially or to complete them all at once—each scenario is independent of the others. Thus, if this professional development module is spread out over multiple days, the scenarios might be an effective way of reviewing previously learned concepts before returning to Parts 5 and 6. The accompanying Scenario Consensus Guide provides a summary of the most salient positive and negative consequences associated with each scenario.

Closing Activity: In Part 2 of the module, a set of common rationales or justifications for why teachers feel compelled to use test-preparation activities is presented. As you proceed through the module, you should gain insight regarding how some of the possible actions taken based on these rationales are problematic. During the closing activity you will be asked to summarize your thoughts and ideas corresponding to these actions. You can then compare your perceptions to those provided in the handout titled *Closing Activity Feedback*.

Part 1: What is “test preparation?”

In a general sense, test-preparation activities are actions that are being taken to promote a student’s readiness to take a test. These actions include teaching students the general content and skill areas covered by the test (e.g., reading comprehension, math problem solving), as well as practicing the skills that are needed for students to demonstrate their command of this content/skill when taking the test (i.e., test-taking skills).

Activities conducted in the name of “test preparation” vary in terms of their appropriateness. Some activities are always OK, some are never OK, and many more can be characterized as—“it depends.” Determining the “appropriateness” is best achieved by considering the negative consequences associated with the results of using the particular activity because many activities appear to be quite appropriate until a closer look reveals the potential harm associated with their use in a given context. For example, some practices that are appropriate for classroom assessments are inappropriate when used in conjunction with the *Iowa Tests*.

Examples of different types of test-preparation activities include:

- Practicing with exactly the same questions that will be on the test to be administered
- Practicing with questions from a previous version of the test
- Targeting instruction to cover only the content areas to be tested
- Reviewing content/skill areas with students before taking the test
- Using primarily multiple-choice questions on all classroom tests
- Having students complete practice tests in advance of testing (e.g., commercially developed/published or teacher made)
- Teaching test-taking skills (e.g., timing, completion of bubble sheets, strategies for answering multiple-choice questions)
- Providing instruction on the general content and skill areas covered by the test (e.g., reading comprehension, math problem solving)

At the completion of this module you should be able to explain why some of these activities are never appropriate and why at least one of these activities might always be appropriate. Furthermore, you should be able to summarize the factors associated with the remaining activities that contribute to whether their use is appropriate in a given context.

But before discussing the appropriateness of certain test-preparation activities, it is helpful to gain a better understanding about why educators feel that these types of activities are important (Part 2 of this module) and why educators should care about the appropriateness of these activities (Part 3 of this module).

Part 2: Why do educators feel that test-preparation activities are necessary and/or important?

To help answer this question, intensive phone interviews were conducted with over 90 Iowa teachers during the spring of 2006. Based on these interviews, it's clear that the vast majority of teachers believe that at least some type of test preparation is necessary in order for test scores to be a true reflection of student learning.

Furthermore, the justifications provided by teachers for using particular types of test-preparation activities were nearly always based on good intentions, as can be seen by the sample of justifications provided below.

"I want to make sure that my students

... do really well on the test(s) and score as high as they can."

... know how to answer multiple-choice questions."

... remember what I taught them earlier in the year."

... have had the opportunity to learn the content and skills covered by the test."

... have the same advantages as those students in Mr. Whosit's class."

... feel comfortable and not anxious when taking the test."

... feel good about themselves and their performance on the test."

... can understand what the questions are about, especially special education students and ELLs."

... are familiar with the terminology or vocabulary used on the test."

... are testwise and know the "tricks" when taking a multiple-choice test."

... receive scores that represent their true knowledge and not just their ability to read the question."

But, are good intentions enough to justify the use of a particular activity?

Of course not. As educators (and possibly as parents), many of us have first-hand experience with considering the consequences of the actions of children, as well as the actions of others, and are fully aware that having good intentions is rarely sufficient to justify an action.

To clarify this point in the context of test preparation, consider the types of actions that might occur in response to an educator’s “good intentions,” as illustrated in the table below. What if the good intention listed on the left leads to the action listed on the right?

Good Intention	Action
... do really well on the test(s) and score as high as they can	→ giving students more time to take the test than what is stated in the <i>Directions</i>
... know how to answer multiple-choice questions	→ using multiple-choice questions almost exclusively on classroom assessments
... remember what I taught them earlier in the year	→ spending two weeks right before the test reviewing and practicing content known to be on the test
... have had the opportunity to learn the content and skills covered by the test	→ conducting a mini lesson right before the test to cover a particular content area on the test that’s taught later in the year
... have the same advantages as those students in Mr. Whosit’s class”	→ using the <u>same</u> materials as those used by Mr. Whosit to “prepare” students
... feel comfortable and not anxious when taking the test	→ using commercially-prepared practice tests the week before the test is administered
... can understand what the questions are about, especially special education students and ELLs	→ reviewing the test questions in advance with these students to help them understand what the questions are asking
... are familiar with the terminology or vocabulary used on the test	→ using last year’s test to practice with the students
... are testwise and know the “tricks” when taking a multiple-choice test	→ teaching test-taking skills for two days right before the test
... receive scores that represent their true knowledge and not just their ability to read the question	→ reading out loud the questions on the <i>Reading Comprehension</i> test

The negative consequences associated with implementing some of these actions are probably more obvious for some actions than for others. By the time you have completed this lesson, you should be able to not only identify the negative consequences associated with each of these actions, but also be able to describe a more appropriate action that could have been taken, if possible, in order to achieve the same goal. During the closing activity you will be asked to summarize your thoughts and ideas corresponding to these actions.

Time for reflection and/or interaction:

What types of things do you do (or use) with students, right before testing or throughout the year, to help prepare them for the Iowa Tests?

Why do you use each of these activities?

Part 3: Why *should* you care about the “appropriateness” of test preparation?

As educators, our decisions are constantly guided by considerations about what would be in the best interests of our students. Decisions regarding test-preparation activities should be no different. Sometimes, however, exactly how students might benefit or be harmed by a particular activity is not all that clear. Thus, it’s also important to consider how you personally might be impacted, or how your school or district could be impacted.

For example, consider the following headlines:

“TEACHER IS ACCUSED OF CHEATING ON STATE TEST”

“SCHOOL UNDER SUSPICION FOR CHEATING”

What if one of these headlines about your school was in your local paper? In the *Des Moines Register*? What would be the potential fallout?

Here are just some of the potential negative consequences:

- Parents/community might lose confidence in the teacher/school, doubting the trustworthiness and sincerity of future actions.
- Students might believe that “cheating” is an appropriate practice.
- Teachers or administrators might be suspended, fired, and/or have their licenses revoked.

Even if the allegations are not true, the simple suggestion of impropriety can be severely damaging to the integrity of the teachers and/or school. More importantly, the damage to students (e.g., belief that cheating is acceptable, or even expected) might be irreparable.

Let’s take a closer look at what “cheating” means. According to *Encarta® World English Dictionary* © 1999 Microsoft Corporation,

“to cheat” means . . .

- to deceive or mislead somebody
- to break the rules in a game, examination, or contest, in an attempt to gain an unfair advantage

We probably all share a similar understanding of what it means for a student to “cheat.” But what does “cheating” look like when it’s done by an educator in an assessment context?

to deceive or mislead somebody:

actions taken by an educator that result in test scores being higher (or lower) than they should be (i.e., misrepresentation of student achievement)

to break the rules in a game, examination, or contest, in an attempt to gain an unfair advantage:

actions that are not consistent with the standard conditions specified for the administration of the test and/or school policy

Although the potential allegation of “cheating” is an obvious reason for caring about the appropriateness of test preparation, hopefully an even more compelling reason is the negative consequences associated with students receiving scores higher than they really should be. This reason should be made more obvious as the three criteria presented in Part 4 are reviewed.

Time for reflection and/or interaction:

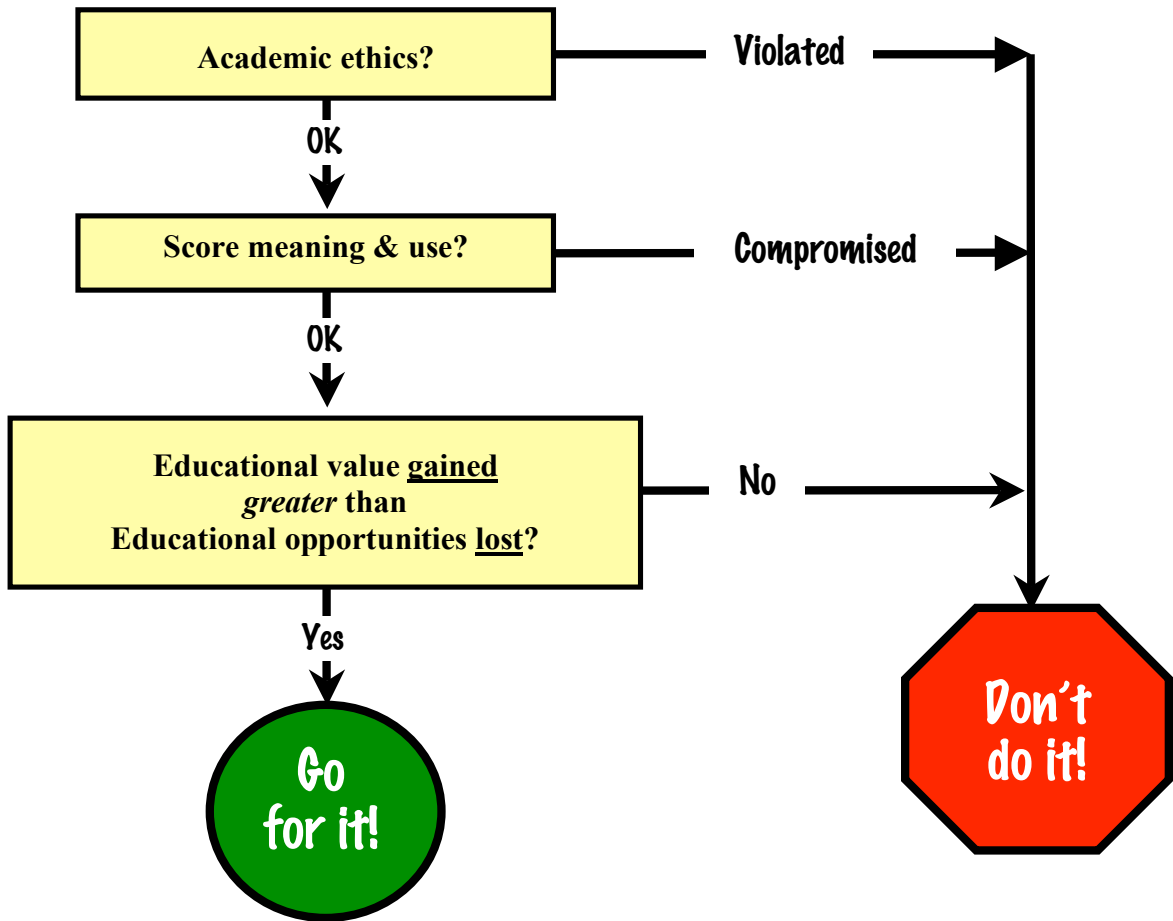
Are there any test-preparation activities that have been used in your school (past or present) that could possibly be considered as “cheating?”

Part 4: How can the appropriateness of a test-preparation activity be determined?

The appropriateness of a given test-preparation activity is not solely determined by whether or not it could be construed as “cheating” or the “misrepresentation of student achievement.” This criterion is actually part of a broader one dealing with academic ethics. In addition to academic ethics, two other criteria should be considered when evaluating the appropriateness of using a particular test-preparation activity—score meaning & use, and educational value.

These three criteria—academic ethics, score meaning & use, and educational value—are interrelated. Thus, if a test-preparation activity fails to meet one of these criteria, in most cases the activity would fail to meet the others as well. But because the “problems” associated with particular test-preparation activities are not always easily identified, it is useful to consider these criteria in a stepwise fashion, starting with the most easily recognizable problems and then moving to the more complex and somewhat fuzzy issues. That is, first determine if the action results in a violation of academic ethics. If there is no apparent violation, then consider if the action might compromise the meaning and use of the scores. If the meaning and use of the scores would not be compromised, then determine if there is any educational value gained in using the activity. If there is educational value, the final consideration is to determine if the educational opportunities being gained by using the activity outweigh those that are lost.

This stepwise process has been summarized in the figure below and is being used as the framework for the remaining portion of this part of the module. These three criteria will be presented one at a time, and will examine the types of factors that should be considered when determining if the criteria have been satisfied. In addition, the potential negative consequences associated with the use of a particular activity that does not adhere to one or more of these three criteria will be identified.



Criterion #1: Academic Ethics

Objectives:

1. The action should not contribute to the misrepresentation or falsification of information [Iowa Administrative Code, Chapter 25, Standard III(e)].
2. The action should not be perceived by students, parents, or the community as being dishonest.
3. The action should not result in a violation of district policy or copyright (e.g., an illegal act).

Standard III of IAC Chapter 25

25.3 (3) Standard III.

... .. *misrepresentation, falsification of information.* Violation of this standard includes:

- e. Falsifying or deliberately misrepresenting or omitting material information regarding the evaluation of students or personnel, including improper administration of any standardized tests, including, *but not limited to*,
 - changing test answers,
 - providing test answers,
 - copying or teaching identified test items, or
 - using inappropriate accommodations or modifications for such tests.

What does it mean to “misrepresent” a student’s achievement? In the context of the *Iowa Tests*, misrepresentation results from reporting scores that are not an accurate reflection of student learning as it relates to the areas covered by the tests. The examples listed in Standard III are the most obvious types of actions leading to the misrepresentation of student achievement, but there are many other less obvious actions. These less obvious actions can be identified when determining if Criterion #2 has been satisfied (i.e., score meaning & use).

Negative consequences associated with using activities that violate academic ethics include:

- Parents/community might question the integrity of the teacher/school.
- Parents/community might lose confidence in the teacher/school, doubting the trustworthiness and sincerity of future actions.
- Students might start to question the teacher’s/school’s trust in their ability.
- Students might believe that “cheating” is an appropriate practice.
- Teachers or administrators might be suspended, fired, and/or have their license revoked.
- Teacher/school could be sued by the test publisher for violation of copyright.
- School/district is classified by the Department of Education as being “in need of assistance” (i.e., placed on the SINA & DINA lists) because inaccurate scores were reported.

Do these types of things really happen? You bet consider some of these examples taken from newspapers around the nation.

Type of Action	Real Examples from the Headlines
Changing test answers	Obvious: Principal told the teachers to correct student's wrong answers. (<i>Education Week</i> , November 13, 1996)
	Not so obvious: After the allotted time for testing, a teacher told students to fill in answers for questions they had left blank. (<i>St Louis Post-Dispatch</i> , April 30, 2005) <u>Note.</u> By allowing additional time the teacher has given students the opportunity to "answer" questions that would have been scored incorrect due to being left blank.
Providing test answers	Obvious: Teachers prompted students with hand signals and pointed to answers. (<i>St Louis Post-Dispatch</i> , May 24, 2005)
	Not so obvious: Teachers signaled students by tapping them on their shoulders to let them know an answer was wrong. (<i>The Huston Chronicle</i> , May 5, 2005) Principal instructed teachers to encourage children to retry specific questions if the teachers thought the children knew the answer but had missed it on their first try. (<i>St Louis Post-Dispatch</i> , March 21, 2006)
Copying or teaching identified test items	Obvious: Curriculum coordinator improperly kept copies of previous exams, and allowed teachers to copy some of them and use them for practice. The so-called "previous exam" was the same version of the test that was subsequently administered to the students. (<i>Education Week</i> , November 13, 1996)
	Not so obvious: Teachers reviewed tests in advance and tailored their instruction to match specific questions. (<i>Education Week</i> , November 5, 2003) Teacher took notes based on the test administered last year and created worksheets for her pupils for this year's test. She also shared the worksheet with other teachers. Some of these other teachers, not knowing the origin of the questions on the worksheet, alerted the principal to similarities between the worksheets and this year's test. (<i>The Baltimore Sun</i> , March 28, 2006)
Using inappropriate accommodations or modifications	Reading test was read aloud to students, resulting in the performance for these students being treated as "non-proficient," regardless of their scores, due to the use of inappropriate accommodations. (<i>Education Week</i> , October 22, 2003)

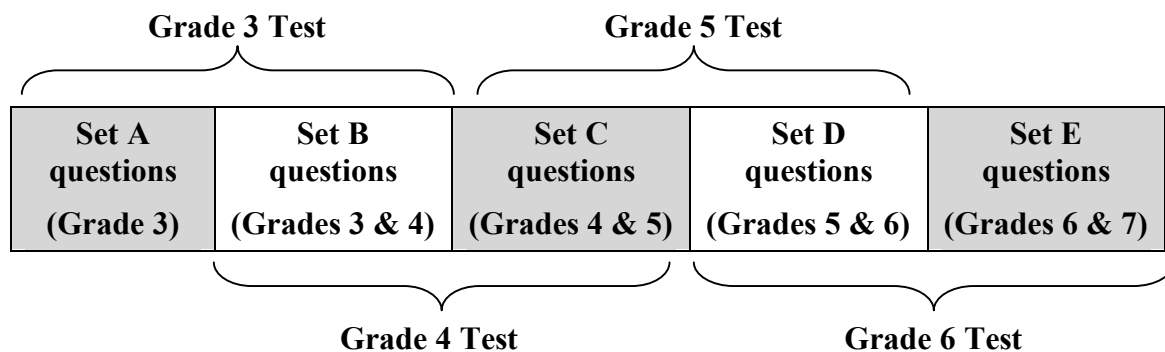
In some of these cases, teachers *unknowingly* used questions from the same test that was to be administered. If you don't know that the questions would be the same, does that make it OK? No, when it comes to the law, ignorance is not an acceptable basis for pardon. Although the intent behind the actions was probably different, the outcome was the same—the scores misrepresent student learning. The educator's lack of knowledge might lighten the sanctions, but it has limited value towards mending the teacher's/school's lost integrity. In addition, due to the fact that the scores no longer are an accurate reflection of student learning, not only has an opportunity to help students in need been lost, but also the school and district will now be treated as a "school in need of assistance" (SINA).

Although teachers should be expected to question the source of materials made available to them for test preparation, additional directives are also needed. One simple solution to making sure that students are not exposed to the test questions in advance of testing is to simply state that no *ITBS/ITED* test materials should ever be used with students prior to them taking the test "for real," or even afterwards. This admonition, however, can carry more weight if educators are aware of particular features of the *Iowa Tests* that can turn good intentions into very negative outcomes.

Important features of the *ITBS/ITED* and how these tests are used in Iowa

1. There are two different versions of the *ITBS* and *ITED* that are currently being used in Iowa—Forms A and B.
2. The test questions on each form of the test never change. The exact same test booklets are used again.
3. These two forms are used alternately in consecutive years. That is, if Form A was administered last year then Form B will be administered this year, and Form A will be administered *again* next year.
4. On each form of the tests there are items that overlap between adjacent grade levels.

The *ITBS* and *ITED* are designed to find out what *all* types of students know and are able to do. Consequently, the collection of tests across grade levels is developed so that some of the exact same questions are asked of students at adjacent grade levels. For example, some of the questions on the tests designed for grade 4 are also on the tests designed for grade 3, and another set of questions on the grade 4 tests will also be on the tests for grade 5. By having these "overlapping" questions it is possible to more accurately distinguish the achievement level of students performing above or below grade level. This overlap is illustrated in the following figure for a given test form. (No questions on Form A are also on Form B, or vice versa.)



So what's the big deal?

Tests are reused every other year. Thus, both Forms A and B are “live” test forms and should never be used for practicing with students.

Even if you think you are using last year's test for practice, it is easy to make mistakes and to use the same exact test that will be administered to your students this year.

If last year's test is used this year for practice (which would be a violation of copyright), the same students will see about half of these exact same questions next year when they take the test at the next grade level. For example, if you used the fourth-grade test from last year (using the above illustration, this test would be comprised of questions from Sets B and C) with this year's fourth graders, next year when these same students take the test in fifth grade they would already have been exposed to about half of the questions—those in Set C. The extent to which this previous exposure assists students in obtaining a higher score next year, contributes to the misrepresentation of achievement.

Let's now turn to a real-life example to illustrate how good intentions and lack of understanding resulted in a very troublesome situation

The Washington Post, February 24, 2002

Amy, the chair of the math department routinely looked at tests in advance to “extract concepts”—to check that she had taught what her students would be tested on. She said that she considered it common practice, a way to make sure her kids had a fair chance to look good.

A few days after the test booklets arrived, the test coordinator (who was also the assistant principal) gave Amy the math portion of the test and told her to “look at them and then lock them away.” Amy subsequently made copies of the questions and distributed them to other teachers during a math department meeting, giving the same advice: “Look them over, then lock them away.”

On the morning of the test, one of the students raised his hand and told the test proctor “I’ve done these questions in math class.”

Upon investigation, it was determined that one of the math teachers had been absent the day that Amy had distributed the test copies and said that he obtained them from another teacher. That colleague had sat through the meeting but said she was not paying close enough attention. Both of these teachers gave the test questions to students for practice, and both said that they thought they were using routine test-preparation materials.

What was the outcome?

Fired: Assistant principal

Suspended: Amy, the teacher who distributed the copies (5 years)
Teachers (two) who used the materials as test preparation (1 year)
Principal (*temporarily*)

Thrown out: Test scores

What is the lesson to be learned?

- Don’t use test-preparation materials for which you cannot determine the legitimacy of the source.
- Don’t make copies of the tests or take notes regarding test questions—for any reason.

What about copyright issues? Does the “fair use” allowance for educational purposes in the copyright law make it OK to use questions from copyrighted tests? No!

The copyright statement for the *Iowa Tests* includes the following guidance:

These tests contain questions that are to be used solely for testing purposes. No test items can be disclosed or used for any other reason. By accepting delivery of or using these tests, the recipient acknowledges responsibility for maintaining such security that is required by professional standards and applicable state and local policies and regulations governing proper use of tests and for complying with federal copyright law which prohibits unauthorized reproduction and use of copyrighted test materials.

If items from any of the *Iowa Tests* are used to prepare students for testing, it is very likely that the students’ scores will no longer be an accurate representation of their achievement. In addition, it is possible that The University of Iowa (the copyright holder) and/or The Riverside Publishing Company (the publisher) might seek damages for copyright infringement. According to the Association for Test Publishers (www.testpublishers.org/copyrightFAQ/htm), the penalties for copyright infringement may include both civil and criminal penalties, with civil remedies

consisting of an award of monetary damages (statutory, up to \$100,000, or actual damages), attorney fees, injunctive relief against future infringement, and the impounding and destruction of copies and equipment used to make the copies.

For example, Educational Testing Service (ETS) sued a former teacher for purportedly distributing “unreleased” forms of the SAT (i.e., forms that were not explicitly made available for public consumption) for “practice” (*Newsday*, April 8, 2004). The outcome of this suit has not been publicized.

Time for reflection and/or interaction:

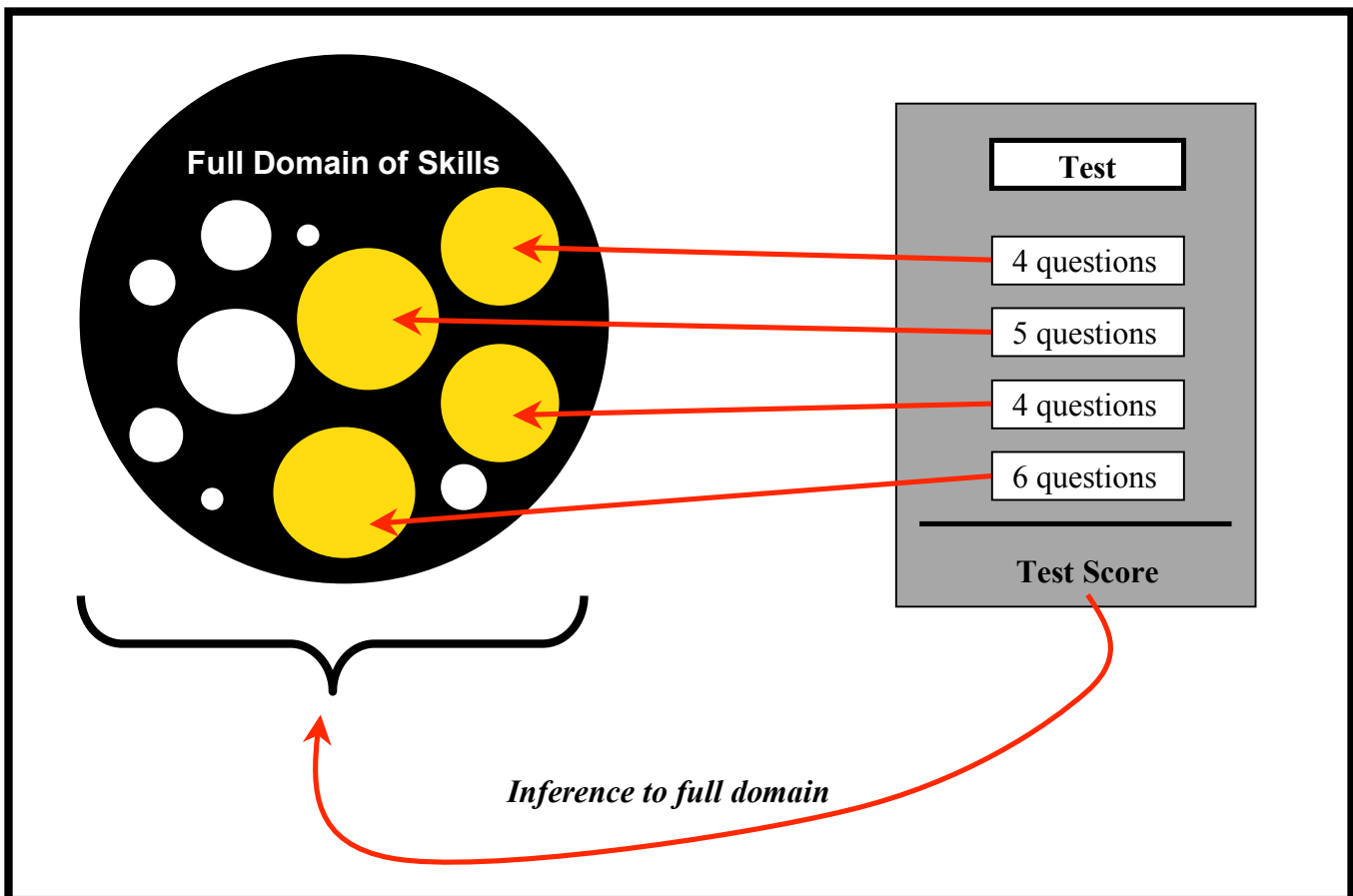
*Do you have (or have access to) any copies of the Iowa Tests (old or current)?
If so, how are you using them? What should be done with them?*

Criterion #2: Score Meaning & Use (also known as “validity”)

Objectives:

1. Test scores should accurately represent student learning related to the specific set of content and skill areas covered by the test.
2. Test scores should not be influenced by a student’s inadequate test-taking skills or limited familiarity with the item formats used on the test.
3. Test scores should allow the user to make an accurate inference regarding student learning related to the larger domain of content and skill areas (i.e., beyond the specific questions on the test).

To help illustrate these three objectives, consider the following figure. The circle on the left represents the full domain (or set) of skills that define a given curricular area the test was designed to measure, such as science. A test that is constructed to measure a student’s attainment of these skills, however, very rarely is able to include questions that completely represent the full domain—due to time constraints and the format of the questions. Instead, a test consists of only a sample of the skills representing this domain, depicted by the shaded circles in this figure. There are often other important skills (depicted by the white circles), such as writing a science lab report, that tests like the *Iowa Tests* are not able to assess. Even for those areas that are covered by the test (i.e., the shaded circles), the actual questions included represent a very small sample of the questions that potentially could have been asked.



The first objective associated with score meaning & use, is to have a student's score be an accurate representation of what the student knows and is able to do in the specific content and skill areas covered by the questions on the test—the shaded circles in this figure. One way to help ensure that this objective can be achieved is to make sure that the student is familiar with the format of the items used on the test, as well as other critical test-taking skills (objective #2). Describing student achievement as it relates to a particular set of questions on a test, however, is not very informative. Instead, nearly always you want to be able to make accurate inferences regarding a student's learning related to the larger, full domain of content and skill areas (objective #3).

Negative consequences associated with having scores that are higher than they should be include the following (in addition to consequences cited for violations of academic ethics):

- lost instructional assistance for students because of inaccurate scores (i.e., students lose out on additional help because their test scores indicate they're doing OK),
- interference with identifying areas of the curriculum/instruction that need improvement,
- inability to use the data to help make correct decisions regarding the effectiveness of a particular type of instructional intervention (if scores are high enough, it's assumed the intervention worked), and
- inability to make meaningful/accurate comparisons across students, classes, or schools (fairness/equity issue) for a given year and/or across time.

Anytime actions taken by a teacher and/or administrator contribute to test scores that do not represent student learning accurately, there is the potential that these actions have directly contributed to the misrepresentation of information. Misrepresentation of student achievement leads to incorrect decision making, and is also considered unethical.

But, isn't test preparation for accountability testing essential so that students will score just as high as they can? According to the guidance provided by Iowa Testing Programs on the development of district policy regarding test use, test preparation, and test security as it relates to the *Iowa Tests* (Iowa Testing Programs, August 2005):

Not really. Obtaining the highest possible test score is not necessarily the primary goal. The notion of test preparation is often associated with students preparing to take a college admissions test such as the ACT or SAT. On such tests, students want to maximize their score to optimize their chance of being admitted or being eligible for scholarship aid. Test-taking strategies that promote the highest possible score are used in conjunction with practice tests to foster greater confidence in anticipation of taking the actual test. However, the purpose of using an achievement test, like the *Iowa Tests*, is to find out just how well a student has achieved. Trying to get the highest possible score, at seemingly any cost, is not consistent with that purpose. There are no particular consequences for the student (no admission decision, for example); the test score should represent what the student knows. Scores that are artificially high are likely to cause some students to get less teacher

attention than is needed and to keep the students from various instructional programs that might help them improve. Thus, scores that misrepresent a student's performance are more likely to be harmful than helpful to that student. Test preparation or test-taking practices that promote artificially high scores could harm students who need extra instructional attention. The use of inappropriate preparation practices may keep a school off the "Watch List" or from being designated as "in need of assistance," but it will not serve the interests of low achieving students whose scores disguise their actual level of achievement. Nor will it serve the needs of the school or district to understand the true achievement of all students so that instructional programs can be modified based on student needs.
(pp. 7-8)

Let's now look at a scenario illustrating the use of practice tests and consider in what ways this practice might result in compromising the meaning of the resulting test scores.

Mrs. Thompson typically uses last year's Advanced Placement (AP) exam to prepare her students for the exam they will be taking in the spring. She also uses ACT practice tests for the same purpose. Therefore, she thought nothing of taking questions from an old *ITED* and then using them to practice certain elementary areas her advanced students hadn't been exposed to for a few years. Prior to using the questions, just to be on the safe side, she modified them so that they were not exactly the same as the originals. For a week prior to the test, Mrs. Thompson used ten of these "modified questions" as warm-up activities. If any of the questions proved to be trouble areas, she conducted mini-review lessons with her students.

To determine how this practice might impact the meaning of the resulting scores, it's very helpful to once again consider the guidance provided by Iowa Testing Programs (August 2005).

To begin, what if Mrs. Thompson decided to use questions from Form A or Form B? That is,

Is it ever appropriate to use the actual test forms (those used in the current or subsequent year) for test preparation?

No, providing students with test items or test answers in advance of the test is highly unethical. Such activity puts the focus on getting particular test questions right rather than on measuring student achievement in the subject area represented by the questions. It is highly rare that a given test item is so important that its content should be learned by all students or taught to students directly. The questions on a test represent only a small sample from all the questions that could be asked when measuring achievement in, say, science or math. When the exact test is the focus of instruction, the test scores lose their meaning, and they portray an achievement result that is dishonest. (p. 5)

What if Mrs. Thompson used test questions from old forms of the *ITED*—forms that are no longer being used in Iowa—without modifying them? Beyond the issue of violating copyright,

Is it ever appropriate to use previous forms of the assessment (e.g., Forms K and L of the *Iowa Tests*) for preparation purposes?

No, the use of previous forms of the accountability tests for practice or preparation purposes also is unethical. Different forms of the same test are designed to be similar in content and skill level so that scores from them can be used relatively interchangeably to estimate growth and improvement. Although the content on a previous form is not exactly the same as that of the current form under use, the similarity is great. Preparation with the previous test form narrows the focus of student learning and restricts the ability of the user to generalize broadly in interpreting the students' scores. (p. 5)

Given Mrs. Thompson didn't use existing *ITED* questions, but modified them before using them—does that make it OK? According to the guidance from Iowa Testing Programs:

Is it ever appropriate to develop practice tests locally that are similar in content or format to the actual test forms currently in use?

No, when practice materials that essentially “clone” the operational test that is in use are developed, the situation is somewhat akin to using a previous form of the test, Limiting the focus of instruction to such materials, even for a brief but intense time period just before testing, creates limits on the generalizability of the test scores. Furthermore, such “practice tests” move the focus of instruction to isolated bits of content instead of the broader skills that should be the target of student learning. (pp. 5-6)

Sometimes the question of how similar a question used for practice can be to a question that is on the *ITBS* or *ITED* (on current or “old” forms) before it is “too similar,” is difficult to answer. The simplest advice that can be given regarding this issue is that if students practice with questions that are modeled after the specific skills and content areas depicted in *ITBS/ITED* test questions, then the practice questions are *too similar* and the use of this practice is *inappropriate*. This type of targeted practice results in the inability to make accurate inferences regarding student learning related to the larger domain of content and skill areas.

Time for reflection and/or interaction:

What was Mrs. Thompson trying to accomplish with her test-preparation activity?

How could she accomplish this purpose in a more acceptable way?

Let's now turn to a scenario illustrating the review of tested content/skills and consider in what ways this practice might result in compromising the meaning of the resulting test scores.

Teachers and administrators at Southwest Elementary are concerned with their students' low reading scores and they are anxious about how well their students will perform when taking the *ITBS* in November. So, they've decided that for the month of October all teachers will spend 10 minutes during the first morning period working on reading passages with students—focusing specifically on inferential types of questions. The passages that have been collected for this practice are from a wide variety of sources, including some that were written by teachers, but none had been taken from the *ITBS*. The questions are almost exclusively in multiple-choice format because the teachers believe that it's important to give their students experience in answering these types of questions.

Is it appropriate to provide students with a review of content covered by the test (in this case, *inferential understanding*) as a form of test preparation? According to the guidance from Iowa Testing Programs (August, 2005):

It depends. A review of content is a common instructional strategy used prior to many forms of classroom assessment. But when the review is narrow and limited to the exact skills that will appear on the accountability assessment, the practice is more questionable. And when such reviews are conducted during the period immediately preceding the administration of the assessment, the practice is unethical. Some forms of review are ethical, but the more closely the focus is on the subskills to be assessed and the more likely the goal is to enhance short-term learning, the more inappropriate the activity would be. Content review geared toward enhancing retention of skills learned previously, however, is a form of sound instructional practice.

The distinction between appropriate and inappropriate subject matter preparation is not always clear. Activities directed towards specific content known to be on the test and conducted shortly before testing time are probably inappropriate. When the purpose is drill for short-term retention, as cramming typically is, the practice is inappropriate. When the purpose is an additional opportunity to review and learn material for which instruction was provided previously, and the focus is on skills that may or may not be covered directly by the upcoming test, the practice is more appropriate. Here are two relevant questions to ask in trying to make the distinction:

- Would the same content-oriented test-preparation activities be used if the current accountability assessment tool were replaced by another that aligns with the district's content standards?

- Would these same content-oriented test-preparation activities be used as scheduled even if the date for administering the assessment were to be moved to two months later?

If the test preparation is designed primarily to fit the accountability assessment tool, or if it needs to be given just before the assessment is scheduled to be given, the activities are probably too narrow in focus and directed too much at short-term effects. They would be considered inappropriate on either basis. (pp. 6-7)

In the Southwest Elementary scenario, teachers were providing focused instruction on inferential understanding of various sources of text. This skill area is a critical component of reading and would likely be included in the district's standards as well as on most tests designed to measure reading comprehension. However, the timing of this additional assistance is suspect. By implementing this focused instruction the month before the test is to be administered, it appears as if the intent is to raise scores rather than to foster the students' long-term retention of this important skill. This is the sort of instruction that would be most beneficial if delivered throughout the year.

What about the fact that the questions were almost exclusively in multiple-choice format? Is this OK? According to the guidance from Iowa Testing Programs (August, 2005):

The appropriateness of any proposed practice should meet either of the two following standards:

- It will promote the learning and retention of important knowledge and content skills that students are expected to learn.
- It will decrease the chance that students will score lower on the test than they should due to inadequate test-taking skills or limited familiarity with the item formats used on the test.

Activities that do not meet one or the other of these criteria are more likely to be unethical, to promote only temporary learning, or to waste instructional time. (p. 4)

If teachers at Southwest Elementary rarely use multiple-choice questions to assess their students' understanding of what they've read, students might be unfamiliar with how best to answer these types of questions and would likely benefit by having *some* practice so that their scores are a more accurate reflection of how well they understand what they have read. An over reliance on multiple-choice questions on classroom assessments, however, can restrict the type of information that teachers can obtain through their assessment process. This type of restriction often results in some important achievement targets—as defined by the standards and benchmarks—not being assessed.

Time for reflection and/or interaction:

What might teachers at Southwest Elementary do more appropriately to build students' inferential understanding?

Criterion #3: Educational Value (gained and lost)

Objectives:

1. The action should promote the learning and long-term retention of important knowledge and content skills that students are expected to learn, as defined by the district's standards/curriculum.
2. The action should provide students with knowledge and skills that have applicability to a broad range of situations/contexts—not just completion of a set of multiple-choice questions.
3. The amount of instructional time dedicated to test preparation should be warranted in light of the types of educational opportunities being replaced/lost.
4. The actions should be matched with the needs of individual students.

Negative consequences associated with lost educational value include the following:

- Student learning is short-term or is lacking in importance.
- Students are not learning all important educational outcomes due to the reallocation of instructional emphasis.
- Students are not being given the opportunity to apply their knowledge and skills to a broad range of situations.

All of these consequences are worse (for students) than if the school makes “the list.”

Let's turn to a scenario illustrating the use of multiple-choice questions on classroom assessments and consider in what ways this practice is related to educational opportunities gained and lost.

Mr. Newton, a 7th grade science teacher has become annoyed and overwhelmed by the continuous edicts from the administration to “get those scores up.” Though he's adamant about not changing the curriculum, he has decided that one logical and beneficial thing to do is to structure most of his classroom tests like the *ITBS* so that students are familiar with the format and language used on the science assessment. To do this, he reviewed the *Interpretive Guide for Teachers and Counselors* and saw that a large number of questions on the Science test were related to scientific inquiry—a skill he teaches but very rarely assesses on his regular classroom tests. Thus, he has made a concerted effort to include multiple-choice questions on his regular classroom tests to measure skills related to scientific inquiry.

It appears as if Mr. Newton has decided to integrate test preparation into his regular instruction instead of right before the *ITBS* is administered. Is this appropriate? According to the guidance from Iowa Testing Programs (August, 2005):

How far in advance of testing should test-preparation activities be used, or for how long prior to the start of testing should each activity be used?

Ideally, test preparation should be an integral part of the regular instructional program rather than an add-on activity. In that way, instructional activities that support test preparation likely would occur throughout the year rather than in a concentrated block of time just prior to the test administration. Activities that occur just prior to the testing tend to have a short-term effect, whether the temporary impact was intended or not. Also, intense practice or attention to testing during the weeks immediately before testing tends to put undue pressure on some students, causing them to be less prepared psychologically for performing at their best. (p. 5)

So, integrating the practice into regular instruction is a good thing. But what about the fact that Mr. Newton has structured most of his classroom tests to be similar to the *ITBS* in terms of item format and language used? What, if anything, has been gained or lost? Do the benefits outweigh the loss?

The emphasis that Mr. Newton has placed on scientific inquiry—a skill directly covered by the *ITBS*—is not problematic because it is a valuable learning outcome as defined by his school's curriculum and because he did not use questions from the *ITBS* as a model of what he should include on his regular tests.

In addition, helping students with the format of the test helps ensure that errors are a result of lack of achievement and not because of a lack of understanding based on how the question was worded. However, if he makes most of his tests resemble the *ITBS* (i.e., nearly all multiple choice questions), he will miss out on other aspects of science achievement, such as use of lab equipment and the written summaries of lab results, which are not easily measured by multiple-choice questions.

As a final scenario, let's look once again at an example of how teaching of test-taking skills might be implemented.

At South Central Junior High, "*ITBS Week*" is a time when everyone comes together in a unified effort to motivate and prepare students for the tests. The counselor has developed a test-taking skills curriculum, and every teacher uses this curriculum with his or her 2nd period students the week before the test for two full class periods. Skills that are practiced include strategies for answering multiple-choice questions, what to do when you get stuck on a problem, and tips for pacing and timing. The curriculum is also designed to motivate students to try their best.

The staff at South Central Junior High are using practices designed to review test-taking skills rather than content that will be on the test; it is encouraging that the efforts are designed to help students do their best on the tests so as to get the most accurate scores possible. However, two full class periods on these skills might be a bit much and probably takes away from other important learning objectives. Additionally, doing the practice right before the test might place undue pressure on some students. This pressure could result in them not being able to do their best, and may result in only short-term learning of these important skills. A few students might benefit from such intense work, but most likely won't. Thus, the activity helps a few and wastes the time of many.

But, what about the fact that all students are being taught this test-taking skills curriculum? Is it important to make sure that all students are treated the same way? Once again, turning to the guidance provided by Iowa Testing Program (August, 2005):

Should all students be provided an opportunity for test preparation in advance of the actual test each year?

Not necessarily. Test-preparation activities probably should be limited to assisting those who need help rather than for use en masse. Younger students may need more help than older ones, lower achieving students may need more than higher achieving ones, and some students in special programs may need more than those in regular programs. Just as general instruction often is individualized, so test preparation (which is a form of instruction) should be individualized. It should be based on need. In addition, the potential negative consequences of test preparation should be considered. Outcomes such as elevated test anxiety or overconfidence about the easiness of the test tasks can result from poorly designed test preparation emphasis. (p. 7)

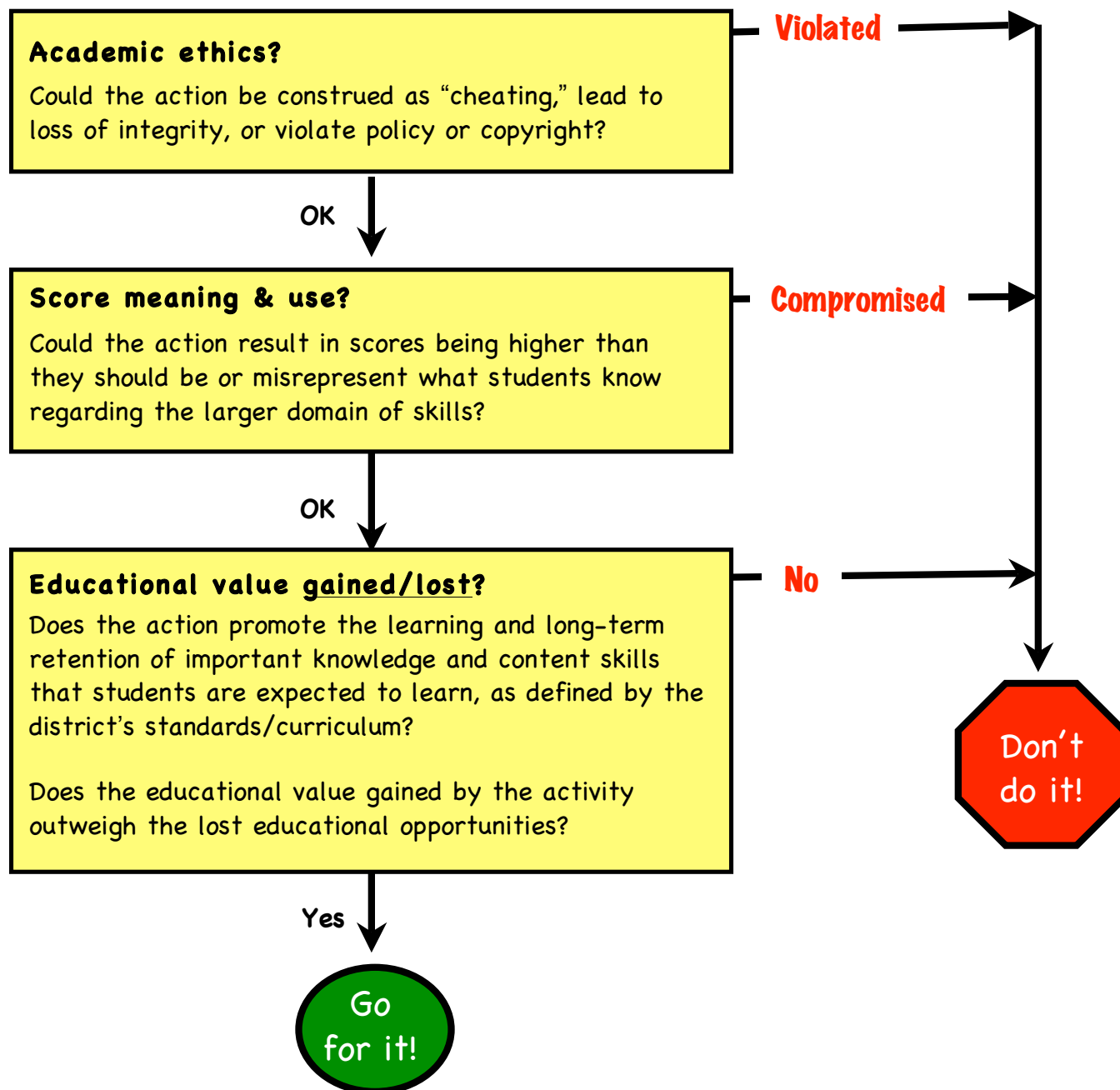
Scenario Activity:

You have now completed Part 4. The examples provided in this part of the module were selected to illustrate some of the most common practices used by teachers in an attempt to prepare students for the *Iowa Tests*. Additional scenarios are available for you to apply what you have learned regarding how the appropriateness of test preparation can be evaluated in terms of academic ethics, score meaning & use, and educational value.

Although you can work through these scenarios on your own, working together in a small group would probably be more beneficial because of the opportunity to interact and exchange ideas and perceptions. There is no need to work through the scenarios sequentially or to complete them all at once—each scenario is independent of the others. Thus, if you are not able to work through the scenarios at this time, you might want to consider using part of them when you return to complete the module as a way of reviewing previously learned concepts before returning to Parts 5 and 6.

The expanded flowchart in the following figure provides a summary of the general questions to be considered for each of the three criteria. (A print-version of this flowchart can be obtained from the “Downloads” page.)

Flowchart: Considering the Appropriateness of Test-Preparation Activities



Part 5 What should educators do when they feel pressured to use inappropriate practices or they are aware of such practices being conducted in their school?

Sometimes, unfortunately, test-preparation practices are used by educators that do not meet the criteria related to academic ethics, score meaning & use, and educational value. In other situations, teachers are being encouraged to use such inappropriate practices. If you were in this position, what would you do? What *should* you do?

In order to work through some of the issues related to this type of situation, let's consider a real-life example—one that was shared when teachers were interviewed in 2006. This is what Zach told us:

Zach: We have teachers that have copies of previous *ITBS* test booklets. They make copies of the tests and use them as practice tests. I think ... this is morally incorrect. I just think it gives teachers a bad name; I don't think it's proving anything as far as what the teachers have actually taught, what the students actually know. Our principal, however, isn't concerned. He doesn't care. He says "I will adhere to ethics when I feel NCLB is ethical."

So, should I say something to someone about this? I mean ... I don't want to get anyone into trouble.... I don't want to get into trouble. What should I do?

What *should* Zach do? The easiest thing, perhaps, would be for Zach to do nothing. After all, this practice is not impacting him directly. In addition, if the scores go up it will make his school look good and, furthermore, going over his principal's head to report this practice to the superintendent is not an action easily taken.

On the other hand, what if Zach's concern is legitimate? What if this practice is a violation of academic ethics? Does he have an *ethical* or *legal* responsibility to report the practice?

According to the Iowa Department of Education and the Iowa Board of Educational Examiners, the answer to this question is a resounding YES. All licensed educators have both an ethical and legal obligation to report test preparation or testing practices that contribute to invalid scores. Although an educator's "fear of retaliation" associated with reporting these types of practices might be real, this fear is not a sound defense for an educator who fails to report testing practices known to be unethical. (Employees are protected against retaliatory actions by employers, by the Iowa Civil Rights Law, Iowa Code Chapter 216.) By reporting such practices, not only are educators taking a professional stand, but they are also putting the interests of students first.

So, what *should* Zach do? It appears as if the academic ethics criterion has been violated by actions of educators in Zach's school. Thus, Zach has both a professional and legal responsibility to report the incident—but, to whom? In situations such as this, it would be

extremely helpful if his district had a policy and/or procedures that could be used to provide guidance on how to handle the situation. The allegations are serious and potentially very damaging to the educators involved as well as to the school, thus great care needs to be taken when investigating the severity of the actions and documenting the evidence.

The guidance provided by Iowa Testing Programs on the development of district policy regarding test use, test preparation, and test security as it relates to the *Iowa Tests* (Iowa Testing Programs, August 2005) was developed, in part, to assist in situations such as this. In addition to district policy, it would also be helpful if specific guidance were available in the form of written procedures. Although it is beyond the scope of this module to delineate the specific procedures that should be followed, which is a district-level decision, the following suggestions are offered. Specifically, the procedures should include answers to the following sets of questions:

1) Who within the district should be notified first of the situation?

- What are the appropriate channels for communication?
- In which situations should the suspicions be reported directly to the Iowa Board of Educational Examiners, and by whom?

2) Upon notification, what type of investigation needs to be made and by whom?

At a minimum, the evidence to be collected should include the following:

- What specifically was done?
- If “previous” forms of the *ITBS/ITED* were used, which ones? Are they “live” forms?
- Which teachers were involved?
- Which classroom(s), grade level(s), and specific students were impacted?
- Which specific tests were impacted (e.g., math, science)?
- How likely is it that the test scores are or will be higher than they should be?

This information is needed to establish the usefulness of the test scores but also to establish facts when consequences are considered.

3) Which agencies should be notified, under what conditions, and by whom?

Anytime that it is believed that something was done related to the *ITBS* or *ITED* that shouldn't have been done—notify Iowa Testing Programs (ITP). ITP personnel can assist you with interpreting the likely impact of the actions taken, and provide advice on steps that should be taken associated with the resulting test scores.

Anytime that it is believed that the resulting scores might no longer be accurate, the Iowa Board of Educational Examiners and the Iowa Department of Education (specifically, the Title I consultant) must be notified.

4) What types of documentation should be made and by whom?

If it is likely that some of the test scores are not accurate, the scores should be flagged. That is, they should be marked or annotated to indicate that the scores were obtained under compromised conditions and, as such, the scores are no longer accurate. This documentation should accompany all reports containing the inaccurate scores so that future users of this information are aware of the inaccuracies.

Time for reflection and/or interaction:

Does your district have a policy or procedures related to test preparation that provides sufficient guidance? If not, what might you do to encourage the development and adoption of such guidance?

Part 6 What can/should be done to increase the likelihood that scores are a true representation of what students know and are able to do in the areas covered by the tests?

By now it should be evident that the use of test preparation should *not* be guided by the desire to increase scores, but rather by the desire to ensure that student scores are an accurate reflection of what they know and are able to do in the areas covered by the tests. When the focus is mainly on the scores instead of the meaningfulness and trustworthiness of the information, it is easy to rely on “quick fixes” that do not contribute to any important learning for the student. In contrast, when the focus is on finding ways of improving student learning, an increase in test scores is likely to follow.

To illustrate, let’s consider two alternative approaches to remedying low performance in spelling. First, let’s listen to Bailey:

Bailey: I’m on a Lead Learning Team at my school. After reviewing our test results we realized that our students didn’t do very well in spelling. So we decided that we needed to look at the *ITBS* test and see what kind of words are listed there. Then we could make sure that we cover those words so that our curriculum lines up with the test.

The response by Bailey’s learning team is a common one, based on the understanding that the things covered by the test should also be things that students are taught. After all, how meaningful would the scores from a weekly spelling test be if students had never been exposed to the words on the test? This close alignment between what’s tested and what’s taught is very important for classroom tests where you usually want to make judgments regarding student performance on a very small set of specific learning objectives. In contrast, as you might recall from the discussion in Part 4 of this module, scores from tests like the *ITBS* & *ITED* should allow us to make accurate inferences regarding student learning related to the larger domain of content and skill areas (i.e., beyond the specific questions on the test). Using the spelling test as an example, one would like to use the scores from the spelling test to generalize about how well students can spell the words on a long unseen list of words. Providing students exposure to the specific words included on the test—before the test is administered—would cause us to over generalize and conclude that students can correctly spell many more words than they probably can.

Now let’s listen to the approach taken in Alex’s school:

Alex: In spelling—our kids have never encountered finding the misspelled word in a series of words. Instead, our weekly spelling tests are based on dictation. So we thought, wait a minute—we need to let them at least have an experience with that. We didn't make this decision just because of the *ITBS*, though. We realized that we are always expecting our kids to look at their writing and be able to tell which words are spelled right and which words are spelled wrong. This is an important skill ... it requires them to be able to think in a different manner. Being more analytical, I guess, being able to analyze the words.

Teachers in Alex's school have taken an approach that is very different from the one used in Bailey's school. Instead of focusing on the specific content of the test items, teachers in Alex's school have considered if there are any characteristics associated with what students were being asked to do that might be unfamiliar, thus interfering with obtaining accurate information about their students' ability to spell. Their consideration, however, did not simply stop at identifying characteristics of the multiple-choice item format used on the *ITBS* and incorporating this format into their classrooms. Instead, they carefully considered the specific skills being measured by this type of item format and related these skills to their expectations of what they want their students to be able to do in order to be "good spellers." In other words, they asked themselves whether this was an important skill.

Previous parts of this module focused primarily on facets of activities that contribute to inappropriate test preparation and have provided suggestions for how some of these activities could be modified to make them more appropriate. To conclude this module, some additional factors that could be considered as "appropriate" test preparation have been summarized. Although these factors might not be considered "test preparation" in the typical way in which this term is used, they are factors that can directly impact the accuracy of the resulting scores. These factors include ensuring that the:

- curriculum is being taught effectively,
- students are ready physically and psychologically,
- students are using appropriate test-taking skills,
- testing environment is conducive to optimal test performance,
- test administrators are knowledgeable and prepared for the task, and
- teaching and learning climate in the classroom and school is positive and productive.

Many of the suggestions that follow are elaborated upon in the *Directions for Administration* that accompanies the *Iowa Tests*. (This is an extremely valuable resource for teachers and administrators.) Other suggestions were formed after learning through interviews about the practices followed by some Iowa schools—practices that do not necessarily contribute to more accurate scores.

Curriculum: “Alignment” vs. “Teaching to the Test”

Scores from the *Iowa Tests* are used to make inferences regarding student achievement related to a portion of the school’s curriculum. As such, the scores reflect student performance as it relates to the educational opportunities provided to the student. It is for this reason that curriculum-test alignment is so important. The extent of alignment or how this so-called “alignment” is implemented, however, can greatly impact the meaningfulness and trustworthiness of the scores. Actions taken to increase alignment often result in increasing the emphasis given to specific concepts “because they’re on the test,” while at the same time de-emphasizing or eliminating other important concepts because they are not covered by the *Iowa Tests*. “Alignment” can also result in providing students exposure to the very specific skills to be tested in advance of testing. Both of these practices are sometimes referred to as “teaching to the test,” and are usually not in the best interests of students.

When exactly does “alignment” result in “teaching to the test?” In this era of accountability, is “teaching to the test” such a bad thing? In considering these questions, let’s see how one teacher, Marcy, has thought about this issue:

Marcy: It would be nice if we didn’t have to worry so much about NCLB, but I do think we’re accountable. I mean, we have our federal guidelines and we may not agree with them politically or professionally, but they are our guidelines and we have to do what we can for the students to help them to succeed within limits... ethical limits.

If our curriculum is aligned with the test and the curriculum is research based to show that kids are learning and kids succeed with it, then I think that yeah, alignment is okay. Then, as a teacher I’m not questioning whether I am covering the right content. I don’t have to worry about, “Am I teaching to the test?” It’s no longer an issue. If you know that your curriculum aligns with the standards for your district, and the district’s standards align with the *Iowa Tests* then that’s good.

“Teaching to the test” is a phrase that often means different things to different educators, but most often it is used to refer to practices that result from having the test drive what is taught in the classroom—something to which nearly every educator is opposed at some level. Because the distinction between curriculum-test “alignment” and “teaching to the test” is often quite blurry, activities undertaken or implemented using curriculum-test alignment as the rationale should be evaluated using the three criteria—academic ethics, score meaning & use, and educational value. If a practice satisfies these three criteria, it’s likely to be educationally sound, resulting in positive learning opportunities for students.

Students are ready physically and psychologically

Understand the purpose for testing: Students who have been told why they are being tested are likely to concentrate harder than those who have no idea how their scores will be used. How much they should be told and how the ideas are communicated depends somewhat on their maturity level.

Comfort/Anxiety level: Students can be highly sensitive and responsive to ideas communicated directly or indirectly by teachers or administrators. Thus, statements such as “This test is going to determine if our school is failing, so you better make sure you do your very best!” are likely to contribute to uneasiness on the part of the students. This uneasiness could translate into performance on the test that was impacted by nervousness, thus resulting in scores that are lower than they should be.

Physically prepared: Ideally, students should be physically prepared for school—every day, not just during *ITBS/ITED* testing. Sometimes, however, it might be worthwhile to remind students and parents of the benefits of getting sufficient amounts of sleep and eating nutritious meals.

Motivation: Most students are proud to demonstrate what they have learned, and special motivational techniques to encourage students to do their best on the tests are *not* needed. When too much emphasis is placed on the tests in an attempt to increase student motivation, negative consequences, such as high anxiety or fear of making mistakes, can occur and result in scores that under represent student achievement. In contrast, when teachers view the tests as being a waste of time or providing information of no or limited use, students are likely to respond by not putting forth their best effort, and also result in scores that are lower than they should be.

As some students progress through middle and high school, their motivation to perform well on the *ITBS/ITED* may diminish somewhat. For these grade levels, students seem more willing to question the importance of the tests and wonder what personal significance the results might have. This questioning is most likely to occur in settings where the students have not been accustomed to receiving test results and meaningful interpretations of their scores in previous years. Thus, a history of feedback regarding test scores is perhaps the best preparation for motivating students to do their best.

In contexts where incentives have been used in an attempt to increase student motivation, schools have often experienced one or more of the following consequences: a) decreased teacher morale due to the belief that their school is more interested in increasing student scores than in overall student learning, b) difficulty in interpreting year-to-year changes in scores due to changes in the context of testing, c) violations of student privacy information (i.e., FERPA) due to public recognition related to some of the incentives, and d) creation of a climate in which students “expect” some sort of compensation for putting forth their best effort on any instructional activity.

Students are using appropriate test-taking skills

There are some types of test-taking skills/strategies that are appropriate for every context and there are other skills/strategies that should only be used for certain types of tests. Thus, before encouraging students to use a particular strategy when taking the *Iowa Tests*, first make sure the strategy is consistent with the purpose and characteristics of the tests.

For example, a strategy that students are often encouraged to use when reading informational texts is to read the questions or headings first before reading the written passage. The questions and headings are to serve as “advanced organizers” and provide students with a purpose for reading. In the context of the *Iowa Tests*, however, reading the questions on the Reading Comprehension test before the written passage can be problematic for several reasons. The primary problem is that although this strategy might be useful when students encounter long passages of nonfiction, it is not all that helpful when students are to read fiction or short nonfiction passages—the types of reading passages included on the *Iowa Tests*. When this strategy is used when taking the *Iowa Tests*, students are more often unable to finish all the questions because they spend much more time reading and re-reading the passage. In addition, this type of strategy is likely to make it more difficult for students to answer more complex questions that require students to make inferences, interpretations, and generalizations.

There are also several so-called “tips” that teachers share with students regarding how to take a test that contains multiple-choice questions. Do any of these “tips” sound familiar?

- If you have no idea what the answer is, choose B or C.
- If the option contain words like “always” or “never” it’s incorrect.
- If the answer is too obvious, it’s probably not the right one because there is some “trick.”

These “tips” are not so much “urban myths” as they are tips for taking “poorly written” multiple-choice questions, and these types of questions are not included on the *Iowa Tests*. Thus, students who attempt to use these types of tips when taking the *Iowa Tests* are focusing their attention on factors that are totally unrelated to the knowledge and skill areas covered by the test and their scores are not likely to be a true representation of their level of achievement—scores are often lower than they should be.

Perhaps the most important test-taking skill from which students of all ages would benefit—on all types of tests—is the ability to use time wisely. Specifically, it would be helpful if students could do the following:

- Begin to work as rapidly as possible
- Set up a schedule for checking progress throughout the test
- Omit questions where no clear answer is obvious and move on to the next question
- Mark omitted questions on the answer sheet so that they can be easily relocated, and so not to get off sequence
- Use the time remaining after completing the test to review and check answers for reasonableness

If students were encouraged to use time-management strategies such as these on regular classroom assessments, it would become quite natural for them to apply these strategies when taking the *Iowa Tests*.

Finally, students should not be encouraged to guess blindly or to use one of the “tricks” mentioned above. Instead, they should be encouraged to use deductive reasoning to eliminate options known to be incorrect and to choose from among the remaining options. Educated guesses resulting from the elimination of at least one of the response options can provide some relevant information about the student’s level of competence, whereas random guessing provides no useful information. Questions answered correctly as a result of random guessing contribute to test scores that over estimate the student’s performance.

Testing environment is conducive to optimal performance

Time of day: The tests should be administered during time periods when students are most alert and attentive.

Number of days: The tests should be spread over several consecutive days instead of being crammed into one or two days. Spreading the testing over more than six days, however, has limited benefit and often increases the need for additional make-up testing.

Number of tests per day: The number of tests to administer per day varies by grade level, and depends on the extent to which students are likely to maintain their focus. At the elementary level, it is advisable to keep the amount of testing per day to about 60 to 75 minutes, whereas high school students can tolerate somewhat longer testing periods (assuming that there are reasonable breaks between tests).

Size of group: It's probably best to administer the tests in the regular classroom settings, in groups less than 30. The physical space that can accommodate large groups of students typically is not well suited to testing because of the small work surface available to students and/or the fact that students are seated too closely together. Larger group sizes also make the distribution and collection of testing materials more time consuming.

Physical environment: The work surface available to each student should be large enough to accommodate both the test booklet and answer document, and there should be sufficient space between students to encourage independent work. Distractions, such as use of the intercom system, should be eliminated during the testing period.

Test administrators are knowledgeable and prepared

Familiarity with students: The tests should always be administered by someone who is familiar with the students and who has a good rapport with them. It is not critical, however, that content area teachers administer their specific subject area tests to the students (e.g., math teachers administer math tests and language teachers administer reading and language tests). In addition to making the organization and distribution of testing materials more difficult and time consuming, this type of administration might actually make it easier for teachers to provide students with "assistance" during testing.

Familiarity with materials and procedures: Teachers administering the tests, and other individuals assisting with the test administration, should be familiar with the materials and procedures. Students can be easily distracted, and if it becomes apparent to them that their teacher doesn't know what to do, it will be more difficult for them to concentrate on doing their best.

Standard directions for administration: When administering the tests, the standard print directions should be followed carefully. It should never be assumed that the students know what to do and how to do it. It is not critical, however, that all students hear the directions all at the same time from the same person, as what happens in some schools where the directions are read to students via the intercom system in order to ensure that each child heard the same set of directions. This practice is *not* recommended because it removes the personal connection students have with the teacher administering the test. In addition, students who have questions

about how to proceed are less inclined to ask or less able to interrupt the speaker, and the speaker is unable to monitor if students are paying attention to the instructions.

Teaching and learning climate is positive and productive

It has been shown that one of the most significant—and negative—impacts of NCLB on Iowa schools has been a decrease in teacher and administrator morale. When teachers and administrators are not happy or do not feel a sense of professional pride in their daily activities, it is difficult for them to maintain a positive, productive learning environment for students. Let's listen to Marissa, and see what can be learned from her experience.

Marissa: Two years ago the principal called us in when the scores came back, kind of sat us all down—you know a little celebration. Last year, my fifth grade performed not as well. We had low students that year. We just got pretty much ripped apart ... "This will not happen again. And whatever you did this year, don't do it again. You'd better find something that works better next time."

So, I decided to model problems like—you know—I took a few of the tougher problems that I remembered from last year and kind of twisted those and did some practice problems the week ahead of time. By doing this, it gave students a good understanding of what types of things they would see without giving them the exact question. I also think it helped with the test anxiety a little bit because it gave them some preparation ahead of time so that they knew what to expect. It's a stressful week. We have kids in tears sometimes. Students definitely get very apprehensive—you know. They know it's a big important two weeks so anything that we can do so they know it's an important test—to alleviate their anxiety—helps everybody.

Clearly the approach that Marissa's principal took was not the most positive nor productive approach. Even if the principal didn't really "rip them apart" or threaten them, this was the way in which Marissa internalized the principal's comments. As a result, Marissa resorted to taking existing *ITBS* test questions and used them with her students—albeit with some changes to the wording. This type of practice will likely result in scores being higher than they should have been, thus being a misrepresentation of student achievement. Furthermore, Marissa recognizes that her students were extremely anxious about taking the tests. In all probability, the anxiety felt by her students is directly related to the anxiety that Marissa felt herself. Students are very perceptive and observant of the mannerisms of their teachers. Thus, great care should be taken in deciding how much is said to students regarding the importance of the tests and the manner and tone in which this information is communicated. In addition, when the importance of the test is stressed by teachers to their students—or by administrators to the teachers—the conversation should always take on a positive, encouraging, productive tone.

Now, let's listen to one last teacher and consider how the leadership in this school has approached some of the challenges set forth by NCLB. As you listen to Jacob, you might want to contrast the environment that Jacob is describing to Marissa's, or even to your own.

Jacob: I feel very lucky to be at my school because when No Child Left Behind came around I know a lot of administrators who threw up their hands and were just like this is impossible, this is ridiculous. But our principal came in and was excited about it. He doesn't approach it as, "Well, here's another thing that we have to do." He approaches it as another opportunity for us to get better at our job and to serve our students more effectively. And that carries over into how we deal with the challenges. I know that is one of the biggest reasons that those quick fixes have gone away... we're not in the quick fix business so much anymore, we're more about... this is more for our students' learning—the complete student.

Time for reflection and/or interaction:

What types of initiatives or actions do you think might be taking place in Jacob's school?

What types of initiatives or activities would you like to see taking place in your school?

Do these initiatives and activities put the interests of students first?

Closing

Hopefully this professional development program has:

- assisted you in developing a more complete understanding of the complexities associated with test preparation or reinforced previously held beliefs regarding these practices,
- provided you with the opportunity to think about the consequences associated with your own personal practices—or the practices of others,
- provided a common language to facilitate conversations at your school regarding these practices and ideas,
- assisted you in recognizing the need for district policy and procedures regarding these types of activities, and
- provided you with the support needed to resist pressure to use some of these types of practices in the future.

As you progressed through this module, at times you might have experienced a lack of consensus (if working with others) or possibly even disagreement with some of the recommendations or guidance being presented. Divergence in opinions is not all that surprising given that some of the distinctions between appropriate and inappropriate test-preparation activities are quite blurry—especially those related to the educational value gained and lost by using a particular activity. In these situations, the “appropriateness” is best evaluated by considering all related factors. Often what is at issue is not the “what,” but rather the “how,” “why,” and “to whom.”

Perhaps the most important idea that you can keep in mind in the future as you consider how best to prepare students for testing is that *the use of test preparation should not be guided by the desire to increase scores, but rather by the desire to ensure that student scores are an accurate reflection of what students know and are able to do in the areas covered by the test.* When the focus is on finding ways of improving student learning, increases in test scores are likely to follow.

Closing Activity:

In Part 2, a set of eight sample test-preparation practices that might result from a teacher’s “good intentions” was presented. As a closing activity, print the handout titled *Closing Activity* from the “Downloads” page. Using this handout, consider what the negative consequences associated with each action might be. That is, consider each action in terms of academic ethics, score meaning & use, and educational value. Using the third column of the table you can make note of which criteria are most likely to be violated by use of the action. Then, in the last column, describe a more appropriate action that could have been taken, if possible, in order to achieve the same goal. Once you’ve completed the closing activity, you can check your responses against those provided in the handout titled *Closing Activity Feedback*, which can also be obtained from the “Downloads” page.

Guidance from Iowa Testing Programs (August 15, 2005)



STATE OF IOWA

THOMAS J. VILSACK, GOVERNOR
SALLY J. PEDERSON, LT. GOVERNOR

DEPARTMENT OF EDUCATION
JUDY A. JEFFREY, DIRECTOR

August 15, 2005

Dear School Administrator:

The two documents attached to this announcement were prepared by the Iowa Department of Education and the Iowa Testing Programs to assist districts and schools in establishing administrative rules and procedures for test use, test preparation, test administration, and test security as well as in establishing board-adopted policies in these same areas. The Department encourages districts to adopt policies regarding their testing program so that it is clear to the school community which assessment procedures the district deems acceptable. It is particularly important that teachers and other district staff that are involved in the district's assessment program understand the school board's expectations regarding acceptable practices and the consequences of using inappropriate activities.

The document entitled Guidance for Developing District Policy and Rules on Test Use, Test Preparation, and Test Security for Iowa Tests was developed by Iowa Testing Program staff at the University of Iowa. Staff at the Iowa Department of Education developed the sample board policy document. Together these two documents are intended to help assure parents, members of the community, and policymakers that all Iowa schools are working toward a goal of providing information that truly describes how Iowa students are achieving.

We hope you will disseminate this information to your board and district staff and discuss it with them. We also encourage you to adopt policies governing your testing program that are consistent with the guidance in the attached documents. Questions about the guidance can be directed to Iowa Testing Programs (319) 335-5408.

Sincerely,

A handwritten signature in black ink that reads "Judy Jeffrey".

Judy Jeffrey, Director
Iowa Department of Education

A handwritten signature in black ink that reads "David Frisbie".

David Frisbie, Director
Iowa Statewide Testing Program

Grimes State Office Building - Des Moines, Iowa 50319-0146

PHONE (515) 281-5294 FAX (515) 242-5988

www.state.ia.us/educate

Helping Communities Meet the Learning Needs of All Their Children and Adults

Guidance for Developing District Policy and Rules on Test Use, Test Preparation, and Test Security for the Iowa Tests

Iowa Testing Programs

(August, 2005)

Introduction

The purpose of this document is to offer guidance to Iowa schools and districts regarding the appropriate use of the *Iowa Tests* for both instructional and accountability reporting functions. The document identifies key components that written policy might incorporate, and it offers a rationale for adopting certain practices with respect to test preparation, test administration, and access to test materials. The recommendations offered are intended to be consistent with the *Standards for Educational and Psychological Testing*. Although this guidance is directed toward the use of the *Iowa Tests*, local policy and protocols might be developed to address the use of all assessments.

Each section of this guidance document covers one general aspect of test use for which policy or rules should be in place. Frequently asked questions are posed to highlight key issues, and responses incorporate positions based on state law, best practice, or standards adopted broadly within the profession. Iowa Testing Programs recommends that each district adopt or create policies and rules that are consistent with these positions.

Purposes of written policy. The primary intent of written policy on test use is to help ensure the integrity of the information obtained from the use of educational assessments. If scores become tainted because of inappropriate practices in either preparing students or in administering tests, the meaning of the scores will be distorted and their value for their original purpose will be diminished or lost. Some educators may not be aware of the consequences of their actions, so policy helps to inform and to create awareness of the potential negative impact that certain practices might produce.

Written policy and rules on appropriate use of tests is essential for communicating the expectations of the district to staff members, for outlining the processes to be followed, and for identifying the potential consequences of violating the policy. Thus, written policy and rules need to be distributed to all administrative, teaching, and support staff, and the district or building procedures associated with using testing materials need to be provided to them as well. It would be appropriate for districts to obtain assurances from staff members that they have read the policy and rules and agree to abide by them.

Key components of written policy. A comprehensive policy would cover activities associated with acquiring testing materials and storing them in the district, preparing students for testing prior to the test administration (including practice), administering tests, accessing test materials after testing is completed, and using test reports in either paper or electronic form.

Ideally, procedures for investigating reported violations of policy should be included, as might reference to actions to be taken with those who are found to be in violation of the policy.

For purposes of this document, a test is a printed test booklet, an audiotape recording of a test, a large-print test booklet, or a Braille test booklet. Testing materials also include answer folders and *Directions for Administration* manuals.

A. Acquiring and Storing Testing Materials

In most school districts, administrative responsibility for testing-related functions is delegated to an individual who is readily identifiable as the District Test Coordinator, and in some cases there is subsequent delegation to building test coordinators. Having an identifiable administrative structure for assessment activities is beneficial in defining lines of communication and in providing authoritative sources of information about assessment policy and procedures for staff members who use tests.

A1. Who is authorized within the district to request and receive testing materials from Iowa Testing Programs?

The number of individuals authorized to obtain materials should be limited so that aspects of storage and accessibility can be monitored. But depending on district size and configuration, it may be appropriate to delegate authority to any of several individuals: principal, curriculum coordinator, counselor, ELL coordinator, or TAG coordinator. In addition to regular annual testing, materials might be needed for new students, students being considered for special programming, or professional development with the instructional staff.

A2. For what purposes are school districts authorized to obtain testing materials from Iowa Testing Programs?

Test booklets are needed for annual testing as well as for testing new students who arrive after annual testing is completed. Some students may be tested for special programming selection (e.g., TAG program) during the year. Occasionally test booklets are a helpful reference when members of the teaching staff are interpreting information from selected score reports or conducting alignment-checking activities. Test booklets should be provided to individual staff members only when they have a professional need to use the materials. In addition, tests should not be given to or shown to parents who request to see the tests, except under direct supervision of a district administrator. Test booklets should not be loaned to individuals outside the district staff.

A3. How should testing materials be stored within the district prior to and during testing?

When materials arrive from Iowa Testing Programs for annual testing, they should be stored in a secure area, according to local policy requirements, with access restricted. Procedures will need to be formulated for each building regarding the distribution and collection of materials prior to and during testing. Test booklets should not be readily accessible to teachers, other district staff members, or students other than during actual test administration sessions.

A4. How should testing materials and test results that remain in the district be stored once annual testing is completed?

Provision should be made for storing materials that might be retained once annual testing is completed and testing materials have been returned to Iowa Testing Programs. Tests retained for late testing, professional development work, or use with score reports need to be stored securely with limited access. (See section D below on accessing materials after testing.) Policy or rules should indicate that those who provide unauthorized access to materials are themselves in violation of district policy.

B. Test Preparation and Practice Activities

Test preparation incorporates a broad array of activities that should be intended to prepare students to exhibit their true achievement when they take an assessment. Some activities may be legitimate in helping students understand the assessment process so that the process itself does not interfere with obtaining a true measure of students' achievement. These kinds of activities should represent a good use of instructional time. Some other activities may be completely inappropriate, especially those that are designed to boost test scores without a corresponding lasting increase in actual student achievement. Such activities are regarded as unethical because they distort test scores, rendering them useless for describing true achievement or improvement. The distortions form "falsified" records of student achievement, which can lead to erroneous decisions about students and schools.

One legitimate preparation practice involves ensuring that students are familiar with common test taking strategies that involve topics such as the use of time, the role of guessing, and responding to unfamiliar test item formats. (This kind of preparation should be directed toward a type of test rather than toward a specific test.) As students move through the grades and gain experience with taking the *Iowa Tests*, however, the need for such activities should diminish greatly. Preparation also could involve informing students of the purposes of testing and how scores will be used so that students are sufficiently motivated to perform well. (The *ITBS* and *ITED*

Directions for Administration manuals provide some guidance about appropriate and inappropriate practices.)

The role of policy with respect to test preparation is (a) to identify appropriate activities to carry out with students in preparing them to demonstrate their true achievement, (b) to identify activities that are likely to interfere with desired score interpretations, and (c) to establish boundaries on the kinds of activities that are worth the instructional time they require.

B1. What specific activities are appropriate to use with students in preparation for testing?

Even the most comprehensive policy cannot anticipate all of the activities that might be created for test preparation purposes. Consequently, explicit procedures need to be in place for gaining approval for the use of any test-preparation materials prior to their implementation. (All materials used for test preparation should be retained by the person who used them.) All inappropriate practices should be delineated in the policy, to the extent possible, to communicate specific actions that are deemed in violation and to describe generally the class of inappropriate approaches to preparation.

B2. What specific activities are prohibited from use with students in preparation for testing?

Written policy should indicate that providing students with actual test questions or specific test content prior to testing is unethical and is a violation of district policy. (Currently, these are questions on Forms A and B of the *Iowa Tests*.) In addition, questions from forms of the *Iowa Tests* previously used should not be incorporated in test preparation. (These are questions in Forms K and L of the *Iowa Tests*.)

It may be helpful to distinguish “teaching the test” from “teaching to the test” in concrete terms. For example, teaching a strategy for solving math story problems in general would be an appropriate way to teach *to the* test, and simultaneously, teach to the district’s standards/benchmarks. But practicing solving math story problems that appear on the *Iowa Tests* (or nearly identical problems) would be teaching *the* test. Some educators seem not to understand the implications of using the test items from the current forms of the *Iowa Tests* as instructional material. Some others do not see why prepping with a previous form of the tests is inappropriate or illegitimate practice. The appropriateness of any proposed practice should meet either of the two following standards:

- It will promote the learning and retention of important knowledge and content skills that students are expected to learn.
- It will decrease the chance that students will score lower on the test than they should due to inadequate test-taking skills or limited familiarity with the item formats used on the test.

Activities that do not meet one or the other of these criteria are more likely to be unethical, to promote only temporary learning, or to waste instructional time. To the extent that they involve purchased materials and teacher preparation time, these activities also are probably a waste of financial resources. When ineffective or unnecessary test preparation activities are used, valuable instructional time is lost as well.

B3. How far in advance of testing should test preparation activities be used, or for how long prior to the start of testing should each activity be used?

Ideally, test preparation should be an integral part of the regular instructional program rather than an add-on activity. In that way, instructional activities that support test preparation likely would occur throughout the year rather than in a concentrated block of time just prior to the test administration. Activities that occur just prior to the testing tend to have a short-term effect, whether the temporary impact was intended or not. Also, intense practice or attention to testing during the weeks immediately before testing tends to put undue pressure on some students, causing them to be less prepared psychologically for performing at their best.

B4. Is it ever appropriate to use the actual test forms (those used in the current year or subsequent year) for test preparation?

No, providing students with test items or test answers in advance of the test is highly unethical. Such activity puts the focus on getting particular test questions right rather than on measuring student achievement in the subject area represented by the questions. It is highly rare that a given test item is so important that its content should be learned by all students or taught to students directly. The questions on a test represent only a small sample from all the questions that could be asked when measuring achievement in, say, science or math. When the exact test is the focus of instruction, the test scores lose their meaning, and they portray an achievement result that is dishonest.

B5. Is it ever appropriate to use previous forms of the assessment (e.g., Forms K and L of the *Iowa Tests*) for preparation purposes?

No, the use of previous forms of the accountability tests for practice or preparation purposes also is unethical. Different forms of the same test are designed to be similar in content and skill level so that scores from them can be used relatively interchangeably to estimate growth and improvement. Although the content on a previous form is not exactly the same as that of the current form under use, the similarity is great. Preparation with the previous test form narrows the focus of student learning and restricts the ability of the user to generalize broadly in interpreting the students' scores.

B6. Is it ever appropriate to develop practice tests locally that are similar in content or format to the actual test forms currently in use?

No, when practice materials that essentially “clone” the operational test that is in use are developed, the situation is somewhat akin to using a previous form of the test, as described above in B5. Limiting the focus of instruction to such materials, even for a brief but intense time period just before testing, creates limits on the generalizability of the test scores. Furthermore, such “practice tests” move the focus of instruction to isolated bits of content instead of the broader skills that should be the target of student learning.

B7. Should district staff be required to obtain approval in advance of using test preparation materials and activities?

Yes, the purposes of advance approval are to ensure that unethical practices are not implemented inadvertently and to ensure that students who need assistance will have the right help available in a timely manner. Consequently, a designated staff member—school administrator or testing coordinator—should be given responsibility for reviewing the procedures and materials that any teachers may want to use. (Teachers should ask whether certain practices are acceptable rather than assuming that they are.) Approvals and denials should be documented in case questions arise at a later time about the legitimacy of any activities.

B8. Is it appropriate to provide students with a review of content covered by the test as a form of test preparation?

It depends. A review of content is a common instructional strategy used prior to many forms of classroom assessment. But when the review is narrow and limited to the exact skills that will appear on the accountability assessment, the practice is more questionable. And when such reviews are conducted during the period immediately preceding the administration of the assessment, the practice is unethical. Some forms of review are ethical, but the more closely the focus is on the subskills to be assessed and the more likely the goal is to enhance short-term learning, the more inappropriate the activity would be. Content review geared toward enhancing retention of skills learned previously, however, is a form of sound instructional practice.

The distinction between appropriate and inappropriate subject matter preparation is not always clear. Activities directed toward specific content known to be on the test and conducted shortly before testing time are probably inappropriate. When the purpose is drill for short-term retention, as cramming typically is, the practice is inappropriate. When the purpose is an additional opportunity to review and learn material for which instruction was provided previously, and the focus is on skills that may or may not be covered directly by the upcoming test, the practice is more appropriate. Here are two relevant questions to ask in trying to make the distinction:

- Would the same content-oriented test preparation activities be used if the current accountability assessment tool were replaced by another that aligns with the district’s content standards?

- Would these same content-oriented test preparation activities be used as scheduled even if the date for administering the assessment were to be moved to two months later?

If the test preparation is designed primarily to fit the accountability assessment tool, or if it needs to be given just before the assessment is scheduled to be given, the activities are probably too narrow in content focus and directed too much at short-term effects. They would be considered inappropriate on either basis.

B9. Should all students be provided an opportunity for test preparation in advance of the actual test each year?

Not necessarily. Test preparation activities probably should be limited to assisting those who need help rather than for use en masse. Younger students may need more help than older ones, lower achieving students may need more than higher achieving ones, and some students in special programs may need more than those in regular programs. Just as general instruction often is individualized, so test preparation (which is a form of instruction) should be individualized. It should be based on need. In addition, the potential negative consequences of test preparation should be considered. Outcomes such as elevated test anxiety or overconfidence about the easiness of the test tasks can result from poorly designed test preparation emphases.

B10. Isn't test preparation for accountability testing essential so that students will score just as high as they can?

Not really. Obtaining the highest possible test score is not necessarily the primary goal. The notion of test preparation is often associated with students preparing to take a college admissions test such as the ACT or SAT. On such tests, students want to maximize their score to optimize their chance of being admitted or being eligible for scholarship aid. Test taking strategies that promote the highest possible score are used in conjunction with practice tests to foster greater confidence in anticipation of taking the actual test. However, the purpose of using an achievement test, like the *Iowa Tests*, is to find out just how well a student has achieved. Trying to get the highest possible score, at seemingly any cost, is not consistent with that purpose. There are no particular consequences for the student (no admission decision, for example); the test score should represent what the student knows. Scores that are artificially high are likely to cause some students to get less teacher attention than is needed and to keep the students from various instructional programs that might help them improve. Thus, scores that misrepresent a student's performance are more likely to be harmful than helpful to that student. Test preparation or test-taking practices that promote artificially high scores could harm the students who need extra instructional attention. The use of inappropriate preparation practices may keep a school off the "Watch List" or from being designated as "in need of assistance", but it will not serve the interests of low achieving students whose scores disguise their

actual level of achievement. Nor will it serve the needs of the school or district to understand the true achievement of all students so that instructional programs can be modified based on student needs.

C. Test Administration

The *Directions for Administration* manuals for the *Iowa Tests* provide explicit guidance about preparing for a test administration. They also include directions for administering each test along with instructions that should be read verbatim to provide a standard testing situation for all students, no matter who gives the test, where it is given, or when it is given. District policy should be clear that all tests should be administered exactly as the publisher's manual specifies. The use of accommodations or modifications also is outlined in the *Directions* manual, but each district needs to establish its own policy regarding the use of accommodations, especially by those who do not have an IEP or 504 plan.

C1. What kind of assistance or advice can be provided to students during the test administration?

It should be made explicit to all who give the tests that assistance during the test administration is not permitted (unless accommodations specified in the IEP or 504 plan call for a certain type of assistance). Students should never be told during the test administration which answers are correct or whether they have answered an item incorrectly. The test administrator should not provide hints, word meanings, pronunciations, or rephrased questions.

Expectations for students also should be addressed by the district policy, and these expectations should be communicated to all students. Students should not engage in any practice that could artificially impact the score of any student. Prohibited behaviors include communication with one another in any form during the testing sessions, copying from others, and using electronic devices other than approved calculators or assistive technology.

C2. What are the district's expectations regarding the use of the time limits provided in the *Directions for Administration*?

All time limits given in the *Directions for Administration* should be adhered to strictly (except in the case of an accommodation for an individual student that calls for extended time). Time limits should not be modified at the discretion of the test administrator. Furthermore, the *Directions* for the *Iowa Tests* do not provide for warnings about time remaining or a need to guess at final answers. Thus, these activities are inappropriate.

C3. What are the acceptable and unacceptable accommodations or modifications that might be considered for use with students who have an IEP or 504 plan?

District policy regarding the use of accommodations during testing should be developed in conjunction with policy about including assessment accommodations in IEPs and 504 plans. That is, there should be a list of accommodations from which developers of IEPs or 504 plans might choose, as needed, and a list of changes or modifications that should not be made. For example, a reading comprehension test should not be read to students who struggle with reading. Such a change would modify what is being measured (listening comprehension rather than reading comprehension in this case). Thus, a read-aloud accommodation for reading comprehension should never be permitted and should not be written into an IEP or 504 plan. The use of a dictionary on the Vocabulary test or the use of a calculator on the Math Computation test are additional examples of changes that should not be permitted. The *Directions for Administration* manual for the *Iowa Tests* has additional details about the use of accommodations and modifications.

C4. What procedures should be used in deciding whether a student who has no IEP or 504 plan should be given any accommodations?

Some students who have not been identified for special education services, and thus have no IEP or 504 plan, may have skill levels that require an accommodation. District policy is needed to decide what procedure will be followed in determining whether a request for accommodations can be granted for use with a student who has no plan. The decision should not be left to each individual test administrator. The procedures should include a list of acceptable and unacceptable accommodations, an indication of who decides, and a description of the type of information needed to make a final decision. Each decision, and the procedures leading up to it, should be documented. Decisions should be based on individual student needs and be consistent with accommodations needed and used during instruction. In other words, if a student does not require a certain accommodation during regular instruction, the student probably doesn't require the accommodation during testing either.

C5. Which accommodations are permitted or not permitted with English Language Learners?

The rationale for using accommodations with an English Language Learner (ELL) is the same as the one that applies to a student with a disability. For an ELL, competence in the English language is regarded as a form of disability because the student's low skill level in English interferes with measuring achievement in areas such as science, reading, and math problem solving. To reduce the impact of the language on measuring the student's achievement, changes in the administration might be made. Specific accommodations that could be considered are identified in the *Directions for Administration* for the *Iowa Tests*. Any accommodation that could change the nature of what the test is measuring should not be used. Thus, a read-aloud could be appropriate in math but not for reading comprehension. The use of a translation dictionary would be appropriate in science if meanings of scientific words were not given in the dictionary.

D. Using Test Materials After Testing

There are some legitimate reasons for a school district to have test materials in its possession well after testing has been completed and the bulk of testing materials have been returned to the scoring center. For example, several score reports can be used most effectively if the user has access to the exact test items to which students responded. (The *Group Item Analysis* and the *Class Item Response Record* are both examples.) Curriculum work involving alignment checking is also a situation in which access to the test booklet may be needed. Districts may request copies of tests from Iowa Testing Programs for such uses. (See section A above, also.)

D1. What kinds of alterations to a student's answer document are permissible by individuals other than the student once testing is over?

In preparing answer documents for scoring, it may be appropriate for a staff member to erase smudges or stray marks on student documents. But no changes should ever be made to student responses to test items except by a student during testing. If demographic coding of answer documents needs to be completed by a district employee, explicit directions should be given about how to code and which existing marks should not be changed. Student answer documents should be stored temporarily in a secure area to prevent unauthorized access to them.

D2. Who should be permitted access to testing materials during the year?

Individuals who are authorized to acquire test materials from Iowa Testing Programs should be vested with the responsibility for making tests accessible to staff members who have a professional need to use them. Procedures regarding who has access privileges, the purposes for which access can be granted, and the duration of possession should be developed for policy implementation purposes. It is inappropriate to disclose the contents of test materials to students, parents, or others who do not have a professional need for such information. Thus, sharing correct answers to questions or reviewing test items with students after testing is over are both inappropriate.

D3. What procedures should govern the use of test materials outside of the test administration window (e.g., for what purposes or uses can access be granted, for how long can materials be kept, and can materials leave the building)?

Once annual testing is completed, test materials need to be stored until they can be organized for return to the scoring center. Care must be taken to provide the same type of secure conditions as should have been provided when materials first arrived in the district.

For materials that remain, the custodian of test booklets should maintain a log of booklet use: name, date, intended use, due date for return, and actual date of return. Everyone who is authorized to check out a test booklet needs to

be advised about the rules associated with the professional use of the tests. Legitimate uses involve interpreting information on selected score reports and conducting alignment checks between the district content standards and the assessment content. Typically this work is done by groups rather than by individual staff members. Generally, test booklets should be checked out for a day only, and they should not be removed from district buildings where they are intended to be stored securely.

D4. What test materials may be photocopied and with whose prior approval?

Policy should state explicitly that no test pages may be photocopied for any reason by any district personnel. (These are copyrighted materials for which Iowa Testing Programs does not provide permission for copying.) Other testing materials, such as answer sheets and pages from manuals may be copied based on prior permission from Iowa Testing Programs and authorization from the superintendent or designee.

E. Use of Score Reports and Data Files

Results from the *Iowa Tests* are provided to each district in the form of printed paper reports and electronic data files. District policy should specify who has responsibility for the distribution of reports for internal staff use and for external reporting purposes. In addition, the policy should provide for maintaining the integrity of electronic files, both in their original CD-ROM format and within whatever database (local or AEA) the data may have been imported.

E1. Who should receive, store, and distribute test results in their paper and electronic forms?

Many district staff members have legitimate professional needs for possessing and using various printed score reports. Consequently, written policy should recognize those needs. At the same time, test scores of individuals need to be handled confidentially, consistent with the Family Educational Rights and Privacy Act (FERPA). However, there is good reason to limit the individuals who are permitted to release district-wide test data to external groups. Few district staff are close enough to the testing process to recognize when a given report is a comprehensive, complete, and accurate representation of district-wide or building-wide performance. In addition, many are unable to provide the interpretive guidance that would be required for promoting sound use of the test results. Thus, the external distribution of a district's test results should be limited to those who are prepared to certify their completeness and offer interpretive assistance to the recipient.

E2. To whom may test results be distributed and for what purposes?

Certainly students and their legal guardians have a right to see their individual scores, but those scores should be treated as confidential, consistent with FERPA. Thus, the scores of individuals may not be shared, on paper or in electronic form, with others without written permission. (That permission probably should be granted only by the superintendent.) The scores of grade groups (e.g., averages or other summary data) are subject to the provisions of the open records law, but their release or distribution should be handled by individuals who are capable of providing the type of interpretive assistance that is likely to reduce the chances that the data will be misused or misunderstood.

E3. Who should be permitted access to the electronic data files?

Electronic files can be modified easily to change test scores or demographic information associated with individual students. Access to the original data file should be limited so that the file can be maintained as the primary source against which checks can be made for establishing the integrity of analyses that might have been performed with the data. When the data file is the source for mandated reporting (i.e., to the state), limited access also should be maintained so that final results can be certified by the district readily.

E4. Under what circumstances should an electronic data file be modified, and how should such changes be documented?

Occasionally errors occur in preparing data files containing demographic information about individual students. When such data are joined with test data, there may be a need to make corrections in the demographics so that accurate subgroup reporting can occur. Procedures for modifying the data file to correct such errors should specify who has responsibility for making changes, how such changes will be verified, and how the work will be documented so that future users will be aware of the changes and the reasons for them. Of course, when a data file is modified, no changes should ever be made to test scores or student test-item responses.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing* Washington, DC: AERA.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for Teacher Competence in Educational Assessment of Students* Washington, DC: NCME.

Code of Professional Conduct and Ethics. Iowa administrative code, 282, 25.3(3)e.

Joint Committee on Testing Practices. (2004). *Code of Fair Testing Practices in Education*. Washington, DC: American Psychological Association.

National Council on Measurement in Education. (1995). *Code of Professional Responsibilities in Educational Measurement*. Washington, DC: Author.

Accountability Test Integrity/Test Preparation

The _____ school district is committed to ensuring the integrity of the information obtained from the use of educational assessments. This policy is intended to apply to two assessments in particular; the assessment used to meet the reporting requirement under the No Child Left behind Act and the assessment used to meet the reporting requirements for the Annual Progress Report to the Iowa Department of Education.

The purpose of this policy is to identify procedures that can ensure assessment results are truly representative of the achievement of students in our district. It is also our intent to create awareness of the potential negative impact that inappropriate assessment practices might produce, to outline processes to be followed, and to identify the potential consequences of violating the policy. If test scores become questionable because of inappropriate practices in either preparing students or in administering tests, the meaning of the scores will be distorted and their value for their original purpose will be diminished or lost.

APPOINTMENT OF DISTRICT TEST COORDINATOR

The district shall appoint a District Test Coordinator, who may in turn delegate responsibility for testing-related functions to one or more Building Test Coordinators. The District Test Coordinator is the _____ (superintendent, principal/director of curriculum/counselor/etc.). The District Test Coordinator is responsible for storing materials from Iowa Testing Programs in a secure area with restricted access both prior to and after the testing period.

TEST PREPARATION

As a function of educating students, staff may prepare students for assessments by providing instruction in the content areas to be assessed. Staff may also prepare students for assessments by teaching general test-taking skills that are applicable to any test or test format.

Staff shall not conduct reviews or drills that use actual test items or identical format items of the accountability assessments, use copies of tests from previous years, or review test-specific curriculum content with students at any time.

(over)

ADMINISTRATION OF TESTS

In the administration of standardized tests, it is a violation of test security to do any of the following:

1. Provide inappropriate test preparation such as any of the following:
 - a. Copy, reproduce, or use in any manner any portion of any secure test booklet, for any reason.
 - b. Share an actual test instrument in any form.
 - c. Use test preparation materials or strategies developed specifically for Annual Progress Reporting or the Annual Yearly Progress report.
2. Deviate from the test administration procedures specified in the test examiner's manual.
3. Provide inappropriate assistance to students during the test administration.
4. Make test answers available to students.
5. Change or fill in answers on student answer documents.
6. Provide inaccurate data on student answer documents.
7. Engage in any practice to artificially raise student scores without actually improving underlying student achievement.
8. Participate in, direct, aid, counsel, assist, encourage, or fail to report any of the acts prohibited in this policy.

After testing is completed, test booklets are to be returned according to procedures established by the District Test Coordinator.

CONSEQUENCES OF POLICY VIOLATIONS

If a violation of this policy occurs, as determined by the Superintendent following an investigation of allegations of irregularities, the Superintendent shall determine whether the integrity of the testing program has been jeopardized, whether some or all of the test results are invalidated, and whether a teacher or administrator has violated the Code of Ethics of the Iowa Board of Educational Examiners as found at 282—Iowa Administrative Code chapter 25.

Reports of students cheating on assessments shall be submitted to the building principal for investigation and disciplinary procedures.

A staff member found to have committed testing irregularities shall be subject to discipline in accordance with law and Board policy. If the staff member is a licensee of the Board of Educational Examiners, the Superintendent shall make a timely report to that Board.

If the Superintendent believes that assessment results are invalid, the Superintendent shall make a timely report to the Iowa Department of Education.

Scenario Activity

Instructions

This activity has been developed to give you an opportunity to apply what you have learned about the appropriateness of various test-preparation activities, and should be conducted after the completion of Part 4 of the module. Although you can work through these scenarios individually, working in a small group would give you an opportunity to interact, and to exchange ideas and perceptions. There is no need to work through the scenarios sequentially or to complete them all at once—each scenario is independent of the others. Thus, if this professional development module is spread out over multiple days, the scenarios might be an effective way of reviewing previously learned concepts before returning to Parts 5 and 6.

The scenarios depict the following types of activities in a variety of contexts:

- **Using items from the *ITBS/ITED*** (Scenarios 1-3)
- **Teaching test-taking skills** (Scenarios 4-6)
- **Using practice tests** (Scenario 7)
- **Reviewing content/skill areas** (Scenarios 8-9)

After reading through the scenario for which you have decided to focus your attention, determine if the action might result in a violation of academic ethics, score meaning & use, and/or educational value. The specific questions related to each of these criteria are provided in the boxes below. For each criterion, identify or discuss what you believe to be the most salient positive and negative consequences—the Pros and Cons—associated with the use of the activity in the given context. As you think about the potential outcomes, in addition to considering “what” was done also consider the “how,” “why,” and “to whom.”

Academic ethics?

Could the action be construed as “cheating,” lead to loss of integrity, or violate policy or copyright?

Score meaning & use?

Could the action result in scores being higher than they should be or misrepresent what students know regarding the larger domain of skills?

Educational value gained/lost?

Does the action promote the learning and long-term retention of important knowledge and content skills that students are expected to learn, as defined by the district’s standards/curriculum?

Does the educational value gained by the activity outweigh the lost educational opportunities?

Once you have reflected upon the potential consequences or outcomes associated with the test-preparation activity corresponding to a given scenario, you can compare your perceptions to those presented in the accompanying *Scenario Consensus Guide* before moving on to another scenario. Please note that the *Scenario Consensus Guide* does not identify all the potential positive and negative outcomes—only those deemed to be the most salient. If time is short and you cannot work through all of the scenarios, please try to find the time to read through the *Scenario Consensus Guide* for each scenario, as each one presents a somewhat unique set of issues.

Scenario Activity

Topic: Using Items from the ITBS/ITED

1. Mrs. Potter is very concerned that her 3rd graders will not understand what to do when they see the *ITBS* for the first time. She wants her students to feel confident and not anxious when they take the test so their scores show what they know rather than be a reflection of errors caused by nervousness, such as incorrectly filling out bubble sheets. She finds a copy of the *ITBS* that was administered a few years ago and takes a few of the questions from each section the students will be taking. For about 3 weeks, she uses them as warm-up activities.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Using Items from the ITBS/ITED

2. While students are quietly taking their tests, Mr. Flanders examines the math section. He notices there are several questions on estimation—a skill his students have been struggling with all year. Thinking he’s just stumbled upon one of those “teachable moments,” he jots down a few of the problems and plans to use them in a short lesson next week when the test is already over.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Using Items from the ITBS/ITED

3. The administration and staff at Roy G. Biv Elementary have decided they really need to focus on their non-proficient and special education students. Feeling that these students need more repetition and “real” examples of what to expect on the test, certain teachers have been designated to work with these students using portions of the tests administered last year. The teachers walk the students through how to figure out the right answers and then have students do more of the problems on their own.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Teaching Test-Taking Skills

4. Mrs. Bucket is a firm believer that learning how to take tests benefits students not only on classroom and college entrance assessments but also in life after high school where tests are potentially involved. Using the curricular units planned for the month prior to the administration of the *ITED*, she has developed several lessons to review pacing and multiple-choice strategies with her 9th graders.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Teaching Test-Taking Skills

5. Mr. Eliot believes in the appropriateness of teaching his 6th graders test-taking skills such as pacing and multiple-choice strategies. However, he wants to be sure the content used in the practices he designs for students is also beneficial. He looks to page 55 of the *Iowa Tests Interpretative Guide for Teachers and Counselors* and sees that the Social Studies test contains items on the “Earth’s features”—a topic he covered earlier in the year. He decides to spend some time the week before the test teaching students specific test-taking strategies using passages, graphs, and diagrams about the earth’s features as practice.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Teaching Test-Taking Skills

6. Mr. Emerson knows that many of his students struggle with taking timed tests. They often get stuck on a question, spending so much time on it that they run out of time before they finish all of the questions. He also knows that unless the students qualify for extended time as an accommodation for special needs, it's important that students not have additional time when taking the *Iowa Tests*. So, instead, whenever he administers the *Iowa Tests* he gives students a two-minute warning. He notifies students that there are only 2 minutes remaining and tells them to "Make sure you put something down for every question—if it's left blank, it's counted as being wrong."

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Using Practice Tests

7. The math department at Tornado Alley High School spent a summer developing math practice tests in response to poor scores on the *ITEDs* over the last few years, which culminated in the school being placed on the SINA list. To identify specific areas of weakness, the staff reviewed previous test results and identified those for which their students scored quite a bit lower. In developing the practice tests, teachers did not have access to actual *ITED* tests but rather a list of general skill areas. The test questions developed by the teachers represent the types of content and skills covered in the *Concepts and Problem Solving* test, such as *Data Analysis/Probability/Statistics*. All math teachers are required to administer these practice tests once a month for the three months preceding the January administration of the *ITED*. They then use results from these practice tests to identify areas where students need additional assistance and provide remediation where necessary. The staff believes that these tests help students become more familiar with the format of the test questions as well as help reinforce previously learned content.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Reviewing Content/Skill Areas

8. Mrs. Graham, a 5th grade math teacher, has always taught measurement skills at the end of the year after her students have been exposed to fractions. Because measurement is an area her students do poorly on in October when they take the *ITBS*, she has decided to move her measurement unit from the end of the year to the beginning and spend less time on fractions. She believes that if her students get just get a few more questions right, the scores will go up and her school will look better in the public's eye.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Activity

Topic: Reviewing Content/Skill Areas

9. Mr. Tippy believes that it's most appropriate to continually review learned content throughout the year to help ensure retention rather than doing review in a "one shot deal." However, he knows that reviewing some of the content that he teaches is more important than others because it is on the test. For example, he knows students are more likely to encounter questions on interpreting political cartoons than specific questions regarding the disbandment of the Soviet Union. Therefore, for about a month before the test, he finds a political cartoon in the media and does a daily warm-up with students, walking them through the process of interpreting its meaning.

Criteria	Pros	Cons
Academic Ethics:		
Score Meaning & Use:		
Educational Value:		

Scenario Consensus Guide

Instructions

Once you have reflected upon the potential consequences or outcomes associated with the “test-preparation” activity corresponding to a given scenario, you can compare your perceptions to those presented in this *Consensus Guide*, before moving on to another scenario. Please note that this *Consensus Guide* does not identify all of the potential positive and negative outcomes—only those deemed to be the most salient.

If time is short and you cannot work through all nine of the scenarios, please try to find the time to read through this entire *Consensus Guide*, as each scenario presents a somewhat unique set of issues.

Scenario Consensus Guide

Topic: Using Items from the ITBS/ITED

1. Mrs. Potter is very concerned that her 3rd graders will not understand what to do when they see the *ITBS* for the first time. She wants her students to feel confident and not anxious when they take the test so their scores show what they know rather than be a reflection of errors caused by nervousness, such as incorrectly filling out bubble sheets. She finds a copy of the *ITBS* that was administered a few years ago and takes a few of the questions from each section the students will be taking. For about 3 weeks, she uses them as warm-up activities.

Criteria	Pros	Cons
Academic Ethics:	None!	<p>If the test “used a few years ago” was the same one that was going to be administered this time, student exposure to these questions would result in the misrepresentation, falsification of information (violation of Standard III of IAC Chapter 25).</p> <p>The action could be viewed as being dishonest, thus hurting the integrity of the teacher & school.</p> <p>The use of <u>any</u> <i>ITBS/ITED</i> test materials is a violation of copyright.</p>
Score Meaning & Use:	<p>Reducing anxiety regarding testing helps students to give their best performance possible, thus yielding a more accurate indication of the student’s learning related to the specific set of content and skill areas covered by the test</p>	<p>If the questions were on the test to be administered, students likely would remember the items and their response would be a reflection of memorization rather than achievement.</p> <p>Even if the questions were not the same as the ones on the test to be administered, the teacher’s instruction will be focused explicitly on the sample of <i>ITBS</i> items. This narrowed focus is likely to result in scores that are higher than they should be (i.e., misrepresentation of information).</p>
Educational Value:	None!	<p>Although the teacher has not dedicated very much instructional time to these “warm-up” activities, the teaching of test-taking skills is more effective if it is done throughout the year, integrated with regular instruction.</p> <p>If the teacher’s concern about lack of familiarity with “filling out bubble sheets” is legitimate, her actions do not address this concern. A more appropriate activity would be to show students a similar answer document and have them practice transferring their answers to the sheets with a couple of multiple-choice questions from her own curricular materials.</p>

Scenario Consensus Guide

Topic: Using Items from the ITBS/ITED

2. While students are quietly taking their tests, Mr. Flanders examines the math section. He notices there are several questions on estimation—a skill his students have been struggling with all year. Thinking he’s just stumbled upon one of those “teachable moments,” he jots down a few of the problems and plans to use them in a short lesson next week when the test is already over.

Criteria	Pros	Cons
Academic Ethics:	The ITBS math scores for his students were <u>not</u> compromised because the practice took place after the test was completed	<p>The action could be viewed as being dishonest, thus hurting the integrity of the teacher & school.</p> <p>The use of <u>any</u> ITBS/ITED test materials is a violation of copyright (regardless of when it took place).</p>
Score Meaning & Use:	The teacher is using this information with students after the test has been completed. Thus, the students’ scores on the test were not impacted.	The teacher has incorporated “live” ITBS items into his instructions. If he uses these items for instructional purposes in future years, it is likely that students’ scores will be higher than they should be on this particular set of skills because of the students’ exposure to the exact same items in advance of testing.
Educational Value:	The teacher is focused on an activity for which the objective is genuine reteaching/learning.	It does not appear as if learning would be negatively impacted. However, this practice would violate academic ethics and potentially compromise the meaning of scores in the future.

Scenario Consensus Guide

Topic: Using Items from the ITBS/ITED

3. The administration and staff at Roy G. Biv Elementary have decided they really need to focus on their non-proficient and special education students. Feeling that these students need more repetition and “real” examples of what to expect on the test, certain teachers have been designated to work with these students using portions of the tests administered last year. The teachers walk the students through how to figure out the right answers and then have students do more of the problems on their own.

Criteria	Pros	Cons
Academic Ethics:	None!	<p>The teachers have incorporated “live” ITBS items into their instructions. The form of the tests administered last year will also be administered next year. And, due to the fact that some of the items are used on tests for adjacent grade levels, it is highly likely that students will see some of these same items next year.</p> <p>The action could be viewed as being dishonest, thus hurting the integrity of the teacher & school.</p> <p>The use of <u>any</u> ITBS/ITED test materials is a violation of copyright.</p>
Score Meaning & Use:	None!	<p>By narrowing the focus of instruction for these students on the specific set of skill/content areas tested, the scores are likely to be higher than they should be for these students and the extent to which the scores can be used to make inferences to a larger domain has been limited.</p> <p>The scores resulting from the use of this same test next year are likely to be higher than they should be (see academic ethics).</p>
Educational Value:	<p>The staff at this school is genuinely committed to helping students who struggle to achieve their fullest potential. It’s also encouraging that instructional time is not being taken away from the average and high achieving students who may not need the extra preparation.</p>	<p>The extra focus on test preparation for these students could result in a loss of instructional time in other important content for these students.</p> <p>The scores for these students might be an over estimate of their level of achievement, thus they might miss out on important educational interventions needed to assist their learning (see score meaning and use).</p>

Scenario Consensus Guide

Topic: Teaching Test-Taking Skills

4. Mrs. Bucket is a firm believer that learning how to take tests benefits students not only on classroom and college entrance assessments but also in life after high school where tests are potentially involved. Using the curricular units planned for the month prior to the administration of the *ITED*, she has developed several lessons to review pacing and multiple-choice strategies with her 9th graders.

Criteria	Pros	Cons
Academic Ethics:	It does not appear as if the teacher is using content directly connected to what is on the <i>ITED</i> .	None
Score Meaning & Use:	The teacher is focused on teaching lessons that will transfer to other situations rather than just in improving <i>ITED</i> test scores. Might increase the accuracy of the scores	Nothing obvious. But, it depends on the specific strategies being taught. Answering questions correctly by using "tricks" instead of content-related skills can result in scores being on overestimate of student achievement.
Educational Value:	It is encouraging that the teacher is not interrupting her regular curriculum for instruction in order to work on test-taking skills.	The instruction on test-taking skills is only occurring one month in advance of taking the <i>ITED</i> . To be more effective, these types of skills should be taught throughout the year. If the skills are only emphasized right before the <i>ITBS</i> or <i>ITED</i> , students are not likely to internalize these skills and understand the applicability to other contexts. If the skills were taught well before the <i>ITED</i> is given and merely reviewed just prior to testing, this would represent more sound instructional practice. Using activities with <u>all</u> students is a waste of time for those students who do not need the review.

Scenario Consensus Guide

Topic: Teaching Test-Taking Skills

5. Mr. Eliot believes in the appropriateness of teaching his 6th graders test-taking skills such as pacing and multiple-choice strategies. However, he wants to be sure the content used in the practices he designs for students is also beneficial. He looks to page 55 of the *Iowa Tests Interpretative Guide for Teachers and Counselors* and sees that the Social Studies test contains items on the “Earth’s features”—a topic he covered earlier in the year. He decides to spend some time the week before the test teaching students specific test-taking strategies using passages, graphs, and diagrams about the earth’s features as practice.

Criteria	Pros	Cons
Academic Ethics:	The teacher is not using test questions that are on either the current or previous forms of the <i>ITBS</i> .	None
Score Meaning & Use:	The teacher is teaching test-taking skills in the context of the class he teaches. Contextualizing these skills in this way could help students transfer the skills to other types of assessments.	By spending time right before the test reviewing content to be tested, there is the possibility that student scores will be higher than they should be. Although the teacher’s intent was to assist students with their test-taking skills, by using a content area known to be on the test the scores for these students could be higher than what they would have received if the test-taking strategies had been practiced using a different set of content/skill areas. When this type of practice occurs, students have been given an advantage over other students and it is no longer possible to make fair comparisons across groups of students.
Educational Value:	The teacher is reviewing previously learned content, which is generally an educationally sound practice. But, by using only content related to the Earth’s surface, the teacher is placing greater emphasis on this content area than on other important areas. Reviewing content that has not been taught yet would not be educationally sound (e.g., teaching multiplication before addition/subtraction) and thus, would not be appropriate.	The instruction on test-taking skills is only occurring the week before the test. To be more effective, these types of skills should be taught throughout the year. If the skills are only emphasized right before the <i>ITBS</i> or <i>ITED</i> , students are not likely to internalize these skills and understand the applicability to other contexts. If the skills were taught well before the <i>ITBS</i> is given and merely reviewed just prior to testing, this would represent sound instructional practice.

Scenario Consensus Guide

Topic: Teaching Test-Taking Skills

6. Mr. Emerson knows that many of his students struggle with taking timed tests. They often get stuck on a question, spending so much time on it that they run out of time before they finish all of the questions. He also knows that unless the students qualify for extended time as an accommodation for special needs, it's important that students not have additional time when taking the *Iowa Tests*. So, instead, whenever he administers the *Iowa Tests* he gives students a two-minute warning. He notifies students that there are only 2 minutes remaining and tells them to "Make sure you put something down for every question—if it's left blank, it's counted as being wrong."

Criteria	Pros	Cons
Academic Ethics:	None	This practice violates the standard directions for administration, thus could be considered a violation of academic ethics.
Score Meaning & Use:	The teacher is focusing on time management, an important test-taking skills, instead of providing content-related assistance.	<p>When warned that only a short amount of time remains in the testing period, some students become anxious and are not able to put forth their best effort when attempting to answer the remaining questions.</p> <p>The teacher's advice about not leaving any questions blank is prompting students to randomly guess at the remaining questions. Educated guesses resulting from the elimination of at least one of the response options can provide some relevant information about the student's level of competence, whereas random guessing provides no useful information. Thus, questions answered correctly as a result of random guessing contribute to test scores that overestimate the student's performance.</p>
Educational Value:	None	The teacher is teaching test-taking skills during the test instead of in the context of the class he teaches. Contextualizing these skills in this way could help students transfer the skills to other types of assessments.

Scenario Consensus Guide

Topic: Using Practice Tests

7. The math department at Tornado Alley High School spent a summer developing math practice tests in response to poor scores on the *ITEDs* over the last few years, which culminated in the school being placed on the SINA list. To identify specific areas of weakness, the staff reviewed previous test results and identified those for which their students scored quite a bit lower. In developing the practice tests, teachers did not have access to actual *ITED* tests but rather a list of general skill areas. The test questions developed by the teachers represent the types of content and skills covered in the *Concepts and Problem Solving* test, such as *Data Analysis/Probability/Statistics*. All math teachers are required to administer these practice tests once a month for the three months preceding the January administration of the *ITED*. They then use results from these practice tests to identify areas where students need additional assistance and provide remediation where necessary. The staff believes that these tests help students become more familiar with the format of the test questions as well as help reinforce previously learned content.

Criteria	Pros	Cons
Academic Ethics:	The teachers are not using test questions that are on either current or previous forms of the <i>ITED</i> .	None
Score Meaning & Use:	Teachers have used <i>ITED</i> scores to identify areas of weaknesses and have developed practices that cover general topics rather than exact items found on the <i>ITEDs</i> . Additionally, the tests also help students practice using the format of the test, helping to ensure that errors are a result of lack of achievement and not because of a lack of understanding based on how the question was worded.	Nothing obvious
Educational Value:	The teachers are reviewing previously learned content, which is generally an educationally sound practice.	Test preparation activities probably should be limited to assisting those who need help rather than for use with all students. At the high school level, it's likely that lower achieving students may need this type of assistance more than higher achieving students. Given that the practice test is administered once a month for three months, one needs to consider the "costs" associated with spending this much time away from instruction. Does the amount of learning gained outweigh the lost opportunities?

Scenario Consensus Guide

Topic: Reviewing Content/Skill Areas

8. Mrs. Graham, a 5th grade math teacher, has always taught measurement skills at the end of the year after her students have been exposed to fractions. Because measurement is an area her students do poorly on in October when they take the *ITBS*, she has decided to move her measurement unit from the end of the year to the beginning and spend less time on fractions. She believes that if her students get just get a few more questions right, the scores will go up and her school will look better in the public's eye.

Criteria	Pros	Cons
Academic Ethics:	The teacher is not using test questions that are on either current or previous forms of the <i>ITBS</i> .	None
Score Meaning & Use:	The teacher is covering skills that are part of the school's curriculum	Any time the intent associated with a given test preparation activity is focused more on the improvement of scores instead of the improvement of student learning, the activity has the potential to interfere with making sound inferences regarding student learning. In this case, student performance in the area of fractions is likely to suffer because of time being reallocated to measurement. How does this realignment of emphasis impact the overall math score? What types of inferences could be made about changes in performance?
Educational Value:	None	The sequencing of content/skill areas in the curriculum should be based on a thorough understanding of prerequisite skills related to the discipline area and an understanding of how best to structure the learning environment to build essential links between concepts and ideas. Sometimes it is educationally appropriate to rearrange content and sometimes it is not. In this case, students may have difficulty understanding fractions of units of measurement (i.e., 1/8 of an inch), thus making it difficult for them to learn the skill and ultimately limiting their attainment of the skill.

Scenario Consensus Guide

Topic: Reviewing Content/Skill Areas

9. Mr. Tippy believes that it's most appropriate to continually review learned content throughout the year to help ensure retention rather than doing review in a "one shot deal." However, he knows that reviewing some of the content that he teaches is more important than others because it is on the test. For example, he knows students are more likely to encounter questions on interpreting political cartoons than specific questions regarding the disbandment of the Soviet Union. Therefore, for about a month before the test, he finds a political cartoon in the media and does a daily warm-up with students, walking them through the process of interpreting its meaning.

Criteria	Pros	Cons
Academic Ethics:	The teacher is using real life examples from the media rather than cartoons found on the test. This helps ensure that students can apply what they have learned to any cartoon rather than just a select few found on the test.	None
Score Meaning & Use:	The teacher is using real life examples from the media rather than cartoons found on the test. This helps ensure that students can apply what they have learned to any cartoon rather than just a select few found on the test.	Any time a teacher makes a judgment that one content area is more important than another because it is on the test, the teacher runs the risk of focusing on that content, thus inflating the test scores. This inflation would make it more difficult to make sound generalizations about the students' overall social studies achievement.
Educational Value:	The skill he is reviewing, interpretation of political cartoons, is actually a multifaceted skill that involves the ability to understand propaganda techniques, current events, innuendo, etc. These skills transfer to areas other than social studies.	The review is only occurring the month before the test. To be more effective, these types of skills should be taught throughout the year. If the skills are only emphasized right before the <i>ITBS/ITED</i> , students are not likely to internalize these skills and understand the applicability to other contexts. If the skills were taught well before the <i>ITBS/ITED</i> is given and merely reviewed just prior to testing, this would represent sounder instructional practice.

Closing Activity

Instructions:

In Part 2 of the module, a set of sample test-preparation practices was presented that might result from a teacher’s “good intentions.” As a final activity, working individually or as a small group, consider what negative consequences might be associated with each of the following actions. That is, consider each action in terms of academic ethics, score meaning & use, and educational value. Using the third column of the table make note of which criteria are most likely to be violated by the action by using the following coding scheme: “Y” = violated, “N” = not violated, and “?” = no obvious violation.

Then, in the last column, describe a more appropriate action that could have been taken, if possible, in order to achieve the same goal. Once you’ve completed the closing activity, check your responses against those provided in the handout titled *Closing Activity Feedback*, which can be obtained from the “Downloads” page.

Teacher’s good intentions related to students	Corresponding action taken by teacher	Criteria Violated?	What more appropriate action(s) could have been taken?
... do really well on the test(s) and score as high as they can	giving students more time to take the test than what is stated in the <i>Directions</i>	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	
... know how to answer multiple-choice questions	using multiple-choice questions almost exclusively on her classroom assessments	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	
... remember what I taught them earlier in the year	spending two weeks right before the test reviewing and practicing content known to be on the test	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	

Closing Activity

Teacher's good intentions related to students	Corresponding action taken by teacher	Criteria Violated?	What more appropriate action(s) could have been taken?
... have had the opportunity to learn the content and skills covered by the test	conducting a mini lesson right before the test to cover a particular content area on the test that's taught later in the year	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	
... feel comfortable and not anxious when taking the test	using commercially-prepared practice tests the week before the test is administered	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	
... can understand what the questions are about, especially special education students and ELLs	reviewing the test questions in advance with these students to help them understand what the questions are asking	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	
... are familiar with the terminology or vocabulary used on the test	using last year's test to practice with the students	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	
... are testwise and know the "tricks" when taking a multiple-choice test	teaching test-taking skills for two days right before the test	<input type="checkbox"/> Academic ethics? <input type="checkbox"/> Score meaning & use? <input type="checkbox"/> Educational value?	

Closing Activity Feedback

Instructions:

Once you have completed the *Closing Activity*, check your responses against those provided in this handout.

Teacher's good intentions related to students	Corresponding action taken by teacher	Criteria Violated?	What more appropriate action(s) could have been taken?
... do really well on the test(s) and score as high as they can	giving students more time to take the test than what is stated in the <i>Directions</i>	<u>Y</u> Academic ethics? <u>Y</u> Score meaning & use? <u>Y</u> Educational value?	Make sure that: <ul style="list-style-type: none"> • curriculum is being taught effectively • <u>students</u> are ready physically & psychologically to take the test • <u>testing environment</u> is conducive to optimal test performance • <u>test administrators</u> are knowledgeable & prepared for the task • <u>teaching and learning climate</u> in the classroom and school is positive and productive
... know how to answer multiple-choice questions	using multiple-choice questions almost exclusively on her classroom assessments	<u>N</u> Academic ethics? <u>?</u> Score meaning & use? <u>Y</u> Educational value?	Integrate the use of <u>some</u> multiple-choice questions into regular classroom assessments throughout the year for those skills/concepts best assessed using this format. Use other types of questions for those areas requiring other formats (e.g., constructed response, essay).
... remember what I taught them earlier in the year	spending two weeks right before the test reviewing and practicing content known to be on the test	<u>?</u> Academic ethics? <u>Y</u> Score meaning & use? <u>Y</u> Educational value?	Review previously learned content—content that may or may not be on the <i>Iowa Tests</i> —periodically throughout the year to reinforce key concepts. This review should not <u>just</u> take place during the months leading up to testing, but should continue afterwards as well.

Closing Activity Feedback

Teacher's good intentions related to students	Corresponding action taken by teacher	Criteria Violated?	What more appropriate action(s) could have been taken?
... have had the opportunity to learn the content and skills covered by the test	conducting a mini lesson right before the test to cover a particular content area on the test that's taught later in the year	<p><u> ?</u> Academic ethics?</p> <p><u> Y</u> Score meaning & use?</p> <p><u> Y</u> Educational value?</p>	Teach the curriculum in a sequence that best fits the prerequisite skills related to the discipline area and allows for structuring the learning environment to build essential links between concepts and ideas. Sometimes it is educationally appropriate to rearrange content and sometimes it is not.
... feel comfortable and not anxious when taking the test	using commercially-prepared practice tests the week before the test is administered	<p><u> ?</u> Academic ethics?</p> <p><u> ?</u> Score meaning & use?</p> <p><u> Y</u> Educational value?</p>	<ul style="list-style-type: none"> •Review with students, in a positive tone, the purpose of the test and how their scores will be used. •If students have limited experience with this type of test, spend a <u>small</u> amount of time helping students understand how to respond to multiple-choice questions and how to transfer their responses to a separate answer sheet. •When administering the test, make sure all students know what to do prior to starting the test (as indicated in the <i>Directions for Administration</i>).
... can understand what the questions are about, especially special education students and ELLs	reviewing the test questions in advance with these students to help them understand what the questions are asking	<p><u> Y</u> Academic ethics?</p> <p><u> Y</u> Score meaning & use?</p> <p><u> Y</u> Educational value?</p>	Make sure that students are receiving the types of accommodations and/or modifications that best fit their special needs. In many cases the test questions can be read orally to the student (but <u>never</u> on the <i>Reading Comprehension</i> test). Paraphrasing the question, however, is <u>never</u> appropriate.
... are familiar with the terminology or vocabulary used on the test	using last year's test to practice with the students	<p><u> Y</u> Academic ethics?</p> <p><u> Y</u> Score meaning & use?</p> <p><u> Y</u> Educational value?</p>	The terminology used on the tests is considered "standard" for the corresponding discipline area. Thus, if nonstandard terms are used in the classroom it is advisable to also integrate the use of the standard terminology. It is important that students learn the concepts and can recognize them in a variety of contexts—not just those that are used by their teacher.
... are testwise and know the "tricks" when taking a multiple-choice test	teaching test-taking skills for two days right before the test	<p><u> ?</u> Academic ethics?</p> <p><u> ?</u> Score meaning & use?</p> <p><u> Y</u> Educational value?</p>	Incorporate the teaching and/or reinforcement of test-taking skills—as they relate to a variety of item formats—throughout the entire year. The goal is to promote long-term learning of skills that can be used in a variety of contexts.