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—<u>academic ethics</u>, <u>score meaning & use</u>, and <u>educational value</u>. 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

<u>Understand the purpose for testing</u>: 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.

<u>Comfort/Anxiety level</u>: 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.

<u>Physically prepared</u>: Ideally, students should be physically prepared for school—<u>every</u> 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.

<u>Motivation</u>: 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 <u>unable</u> 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 <u>not</u> 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 <u>not</u> 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?