Teacher Made Tests

Pre-Tests:
Determine what students know before they begin a study of topic. Can provide direction for the shape of a unit by identifying students' interests and gaps or misconceptions in their knowledge.

Post-Tests:
Tests given to students after they have studied a subject

Essay Test Items:
Often an effective way to measure students' ability to organize, analyze, apply, synthesize, and evaluate in their own words information they have learned. Require students to use abstract thinking and provide the freedom to think broadly and to express themselves in different ways. May penalize students who have learned a great deal but have difficulty expressing themselves in standard English.

Objective Test Items:
Objective refers to the way in which responses will be graded, not the nature of the question itself. Objective test items are relatively easy to grade and less susceptible to teacher bias in grading. However they encourage guessing and are often more difficult and time consuming to construct. They afford students no opportunity to demonstrate divergent thinking.

Types of Objective Test Items:

Multiple Choice questions
Matching questions
Alternative response questions (also known as true and false)

(a) Multiple Choice Questions

- The stem should state the question or issue
- All choices should be plausible, related, and, ideally, approximately the same length
- Avoid use of the options "all of the above" and "none of the above", which often confuses students
- Preferably, the stem should be stated in a positive form, avoiding the use of terms such as *not* and *never*

(b) Matching Questions

- The items in each column should be related
- The directions should state clearly what they student is to do and any special limitations on the use of choices (e.g. whether items can be used more than once)

- Thought should be given to how many choices there are (more than the number of responses versus the same number)
- The order of the items in each set of choices should be either alphabetical or random
- The number of items in the shorter column should be less than 10 to avoid students wasting time by searching through long lists

(c) Alternative Response Questions

- Each statement should include a single point or issue
- Avoid statements designed to "trick" or mislead students if they do not read carefully
- Avoid terms such as always, never, all, generally, occasionally, and every, which may signal the correct answer

Making Quality Tests

Does each item measure an important learning outcome?

Is each item type appropriate for the particular learning outcome to be measured?

Does each item present a clearly formulated task?

Is the item stated in simple, clear language?

Is the item free from extraneous clues?

Is the difficulty of the item appropriate for the students who are to be tested?

Is each test item independent and are the items, as a group, free from overlap?

Do the items to be included in the test provide adequate coverage of questioning?

Are your test items free of gender, class, and racial bias?

CHAPTER 6

Creating Good Test Questions

Evaluate this item as a good test question, then read student Raul's think-aloud about it that follows.

Question 13: What is the best way to describe the Renaissance Age?

a. all of the below except "d"
b. a period in which all the great artists lived
c. an age of widespread feudalism and rampant religious "correctness"
d. an age that turned scientific and artistic pursuits toward mankind instead of the church
e. an age of rebirth
f. none of the above

Raul's thinking as he responds:

It could be "d" and "e," but also "b," but isn't that just my opinion, not really a fact? Am I supposed to circle the one with the most correct information? Maybe there's one word that's incorrect, and my teacher wants to see if we're smart enough to catch it. Wait, it can't be "b" because other great artists lived in other time periods. Now we're getting somewhere . . . Uh oh, wasn't there worry over "correctness" in the Renaissance as well as the Middle Ages? Okay. Skip this one for now, and see if answering some other question might give me a clue to answering this one . . .

From: Fair Isn't Always Equal
by Rick Wormeli
Stenhouse Publishers
Portland Maine
2006

Such multiple-choice questions as the one just listed have no life in a successfully differentiated classroom. They do not assess students' knowledge and skill. They assess the extent to which students can guess what's on the teacher's mind and respond in a way the teacher thinks they should. Tests shouldn't play games with students' success. Every test question should be important enough to ask and clear enough to answer.

There are many successful test questions that don't waste students' minds or time while also helping us determine a student's mastery, and some of these are even efficient to grade. Let's take a look at what goes into making successful questions.

Use a Variety of Questions/Prompts

Mix traditional and not-so-traditional questions and prompts. Traditional items include: matching, true/false, fill in the missing word, multiple choice, definition, essay, and short answer. Not-so-traditional items include: analogies; drawings; diagrams; analyzing real-life applications; critiquing others' performances or responses; demonstration/performance; integrating more than one topic; exclusion brainstorming; and deciphering content clues that, when put together, reveal a secret message or conclusion.

In addition, we want to mix "forced choice" items with "constructed response" items. Forced choice items are questions and prompts that require students to choose from responses provided by the teacher, such as true/false, matching, and multiple choice. The student does not need to generate the information himself or herself. Constructed response items are questions and prompts for which students must generate the information themselves and apply it in the manner in which it is requested. Examples of constructed response items include opportunities to interpret graphs; write short essays; write short answers; do drawings; or make up analogies, mindmaps, or flowcharts.

By using a variety of questions and prompts, we get a better picture of students' mastery. Some of them will be able to reveal what they know through one format very well, while other students will shine through another. If assessments are supposed to help us get accurate information about mastery so we can adjust instruction accordingly, we want to give students every chance to provide valid renderings of their proficiency.

We can turn more traditional test questions into innovative versions as our students' needs warrant. For example, "Define the Latin word root *terra*" is a traditional test prompt. To push students further, try this: "In the spaces that follow the prompts, write what you think each real or nonsense word basically means. As long as you capture the essence of the root words, the answer will be correct."

Terratempo— _____
Zotox— _____
Noveloc— _____
Lithjector— _____
Sophipsychia— _____
Thermalmaria— _____
Photophobia— _____
Protophytop— _____
Patripathy— _____
Magnijuris— _____

Or students can be given combinations of several common nouns and asked to "coin new words for each combination that incorporates all the nouns listed, using Latin roots and prefixes only."

Include items for which students must generate or purposefully manipulate information. Simply reporting what has been memorized isn't always a sign of understanding and long-term retention, which are our goals. It's easy to parrot information; it's masterful to apply, critique, evaluate, or create it.

Make It Efficient for Students

Provide a "T" or an "F" for students to circle on true/false questions. This way there will be no questions about how to interpret sloppily formed T's and F's. Students' true intentions are clear, and it's not as tiring as writing out the full words may be for some handwriting-challenged students.

For matching activities, write the definitions on the left and list the words to choose from on the right. This way, students read the sentence-length definitions on the left and then scan only the single-word lists to find the correct response, not whole sentences of definitions. It's tiring to first identify a single vocabulary term, then read every single sentence in a long list of definitions, especially if you have a learning disability in reading. Tired students don't produce accurate test results, so let's do everything we can to keep them from getting tired.

It's also helpful to keep matching items on the same page. Flipping pages back and forth gets confusing. Mistakes happen. In addition, keep matching portions of tests to about eight items or less. Beyond eight, it becomes a bit of an endurance test; and once again, it can become confusing and more of a clerical exercise than a thoughtful task that reveals students' mastery.

Nolen and Taylor advised teachers to keep the blanks in fill-in-the-blank items close to the end of the sentence or stem. This prevents reading comprehension issues. In addition, they say that any omitted words that students have to figure out, such as we might use in a cloze or fill-in-the-blank exercise, should

be significant (2005, p. 221). Otherwise, it's too confusing, the answer can go in too many directions, and we won't be assessing what we think we're assessing. They add that it's wise to ask students to explain multiple-choice responses in short answer follow-up questions right after the question (Nolen and Taylor 2005, p. 206). Responses to such questions reveal mastery and non-mastery.

Make sure to highlight key words, such as *three*, *most*, *least*, and *not*, so students don't lose sight of the expectation while forming a response. This isn't making it easier; it's making sure the student reveals what he or she knows.

Double Recording of Test Responses

If you're using a multiple-choice, true/false, one-word answer, fill-in-the-blank, or matching format, ask students to fold their answer papers in half vertically and number the lines exactly the same on both sides of the fold. As they respond to the prompts, they record the answer in the same location on both sides of the fold. For example, if "86.2" is the answer to number 4 on the test, they record that answer on the number 4 blank on each side of the vertical fold (see Figure 6.1).

When students finish the test, they cut or tear the paper down that fold and turn one half in to the teacher. They keep the other half. When everyone is finished, the teacher reviews the answers to the test while students reference their copies of the answers. Students get immediate feedback on how

Figure 6.1 Example of a Double-Recording Answer Sheet

Name:	Name:
Date:	Date:
Period:	Period:
1. _____	1. _____
2. _____	2. _____
3. _____	3. _____
4. _____	4. _____
5. _____	5. _____
. . .	. . .
20. _____	20. _____

they did on the test instead of waiting until the teacher has graded it. Using this approach means students in earlier classes carry around the test answers the rest of the day, of course, which means we'd have to change the sequence of questions for multiple classes; but it's worth doing in order to be timely with our feedback, and it's fairly easy to do in our computer age.

By the way, secondary English teacher, Susan Clark recommends: "When students grade their own papers, ask them to use highlighters or markers so there is no temptation to quickly change answers."

Avoid Confusing Negatives

In general, when assessing students in fifth through tenth grade, we should avoid using response choices that are meant to make students stumble over wording or logic: "All of the above except C and E," "Which of these is NOT associated with . . . ," and "None of these." At those grade levels, such questions don't assess students' mastery. Errors on these items are related more to reading, logical thinking, and worrisome nerves than students' understanding of content. In the last two years of high school, however, dealing with such negative responses is less confusing and can reveal accurate information about our students' understanding of topics. It's okay to incorporate a few of them on tests. Be judicious in their use, however. It's respectful and ethical to remove any question that is unjustifiably complex, used only to see whether students are reading directions or able to think in a contorted manner. Straightforward questions are respectful and useful.

Make Prompts Clear

In his book, *Test Better, Teach Better,* Dr. Popham says, that the less students have to guess the more they can achieve (2003, p. 94). He's correct. If it's a "guess what's on the teacher's mind" test prompt, the assessment becomes a nightmare, and any grade earned is close to meaningless.

"Describe the Renaissance" is an inappropriate essay question. Students don't know where to go with their responses; they don't know what is expected. Truly, the teacher that assigns such a prompt has no one to blame but himself or herself if the student fails in his response.

The effective teacher provides the intended parameters, clarifying for students what is expected. These parameters may include, but are not limited to: a clear example of what's expected, a suggested number of examples that must be included to support the student's claims, approximate length of the essay or project, and a suggested amount of time needed to complete the task. The teacher may want to include the relative point values of every

component to be assessed so that the student knows where to spend most of his or her energy. Based on Popham's urging, here's the Renaissance prompt revised for clarity.

> In 250 to 400 words, describe the rise of intellectual life during the Renaissance. Include in your discussion a brief statement of the impact of any *five* of the following events and people:
>
> - Translating the Bible into English
> - The development of the Gutenberg press
> - Leonardo da Vinci or any one of the inventors/artists of the period
> - Shakespeare, Cervantes, or any one of the author/poets of the period
> - The works of any one of the humanist philosophers (Machiavelli and Thomas More, among others)
> - The Reformation
> - European exploration and expansion to the rest of the world by any one of these: Cortez, Magellan, Pizarro, the Mayflower
>
> This essay is worth thirty points. Each of the five aspects whose impact on intellectual life you describe successfully is worth five points. The remaining five points will be earned by following proper essay format, including a well-crafted introduction and conclusion. This should take no more than forty-five minutes.

In writing our prompts, however, we also need to make sure we don't give away the answer, as in multiple-choice questions that have grammar clues in the stem. For example, using grammar knowledge alone, not our knowledge of landforms, we can correctly answer this test prompt:

> The picture above depicts an example of an:
> A. peninsula
> B. guyot
> C. plateau
> D. estuary

"An" goes with the starting vowel sound in "estuary." If students knew this, they wouldn't need to think twice about their response. If this happens, stop the stem a word earlier, and place the articles in the potential responses:

> The picture above depicts an example of:
> A. a peninsula
> B. a guyot
> C. a plateau
> D. an estuary

Popham wisely points out that some teachers put too many factors into their true/false statements and students don't know which part is the intended response portion. Here's an example:

> True or False: We are able to breathe on Earth because plants produce oxygen and we exhale carbon dioxide.

This is sort of true, but not completely. We get some oxygen from other sources on Earth, and our capacity to breathe has more factors than just the presence of oxygen (like pressure, other gases present, our anatomy, to name a few). A better prompt to assess students' knowledge is:

> True or False: The only factor affecting our ability to breathe on Earth is the abundance of oxygen-producing plants located here.

This version of the prompt removes distracting information. It focuses the student on the one factor the teacher wanted to assess. The score of this test item will yield useful information. Make true/false statements completely true or completely false.

Keep It Short

Two or three will do. We don't need ten similar questions when we can see in two questions whether or not a student understands the concept. It's not a perseverance test. If there are subtle differences that must be assessed, include enough problems or prompts to assess students' proficiency accurately, of course, but less is usually more. If we want to know whether students know how to plot points on a four-quadrant graph, for example, we give them enough ordered pairs (coordinates) to land one in each quadrant, plus a few that place points along both axes, just to make sure they really understand the concepts. We don't give them twenty.

Be Careful of Timed Tests

Author and assessment expert Ken O'Connor reminds us that, "Timed tests are great underminers." He explains that ". . . no one professionally would ever try to collapse their knowledge into one hour of intense performance" (2002). He's right. The idea of timed performances or tests of mastery is a construct of schools, not the working world. This is not to say we shouldn't teach students to be efficient and expedient, but more times than not, we are

assessing students on the extent of their skill and knowledge development, not how much they can cram into a small sampling in a narrow window of opportunity at this early hour on a Tuesday morning in late April. On the few occasions we're assessing fluency or automaticity with ideas, timed tests may be useful, but even then, the result may be inaccurate because students' angst regarding the approaching sound of the buzzer may negatively impact their thinking. It's worth giving serious consideration as to whether time restrictions on tests enhance or impede those tests' ability to reveal what we are seeking about students' levels of proficiency. In most cases, the restrictions impede accurate data collection.

Include Common Errors as Candidates for Responses

Including common errors in responses from which students choose an answer increases the validity of the grade. Students really know their material if they can discern the differences. For example, the answer to a science question could be "rotation," but "revolution" is also on the list of possible choices because the two are commonly confused. The word "weight" could be substituted where "mass" is the correct term to see whether students catch it. Other examples include math answers that vary from one another by one place value and graphs with multiple misinterpretations in the mix of possible responses. In matching responses, provide more choices than questions, and include a few that are similar to one another.

We're not being sneaky by doing this. Spend time ahead of the test explaining that such problems will be on it, and give students ample opportunity to practice spotting subtle errors and unreasonable answers prior to taking the test.

Put Some Fun into Test Questions

Incorporate students' names and their cultures into the test items. Instead of "If a community playground needs enough small gravel to fill a swing set area with the dimensions 40′ × 65′ × 1′, how many cubic feet of gravel will this require?" how about, "Abdul is building a rectangular, practice hockey rink for his championship-winning, Mighty Anoles hockey team. How much water must he pour into the containing walls and then freeze, if the solid ice is 1.5 times the volume of the liquid water, and the dimensions are 100′ × 50′ × 2′?" Students will look forward to the test just to see their names in print or the occasionally outlandish tasks they are doing in its scenarios.

Offering a pun, sharing a topic-related riddle, doing a parody related to the topic is okay once in a while. Humor relaxes students, even if they moan and groan over the puns. Here are some example comments used on an anatomy test: "Did you find the humerus in this test-erus?" or "This is just the tibia the iceberg," and "Grades will be announced to-marrow."

Even with humor and using students' names and their cultures, we never stray from substance. We keep our testing goals in mind as we make such insertions. Instead of "Describe the main character of the novel," we pump it up with, "Create the lyrics to two verses of an [insert name of popular rock star] song that accurately portray what the main character is feeling during this chapter." Instead of "For what did Frederick Douglas fight?" how about, "Give two similarities and two differences between the civil rights policies of our current president and the principles put forth by Frederick Douglas." Believe it or not, students appreciate meaty tasks more than drudgery tasks. Just make sure they have had plenty of practice with similar prompts prior to the test.

Make Sure Questions Assess What You Want to Assess

Sometimes we get so creative and complex that we stray from our goals. To start designing your questions, go back to the essential understandings or questions you've established for the unit of study and design ones that elicit substantive responses to those understandings. There's no need to be tricky; cut to the chase and ask students exactly what you're trying to teach. Here's an example:

Objective: The student will be able to state the difference between osmosis and diffusion clearly.
Test Prompt: What is the difference between osmosis and diffusion?

Straightforward questions often serve us best. Sure we can increase complexity and the compelling nature of test questions by changing the verb as mentioned earlier, but it's always important to be clear about what we're assessing and to get accurate information about a student's understanding. If an interesting new verb or prompt elicits a clear, accurate rendering of mastery, use it. If not, still use it to see students stretching themselves with the topic, but also ask that straightforward question in another prompt.

Make Questions Authentic to the Instruction

If we teach a procedure or concept one way, we test that way. We don't call for an approach on a test question that wasn't practiced by students extensively

during our lessons. Our test questions should be reliable and valid indicators of what was experienced by the students.

If we allow students to use calculators while they practice math problems, we should allow them to use calculators on the test as well. If students are taught to use the writing process in their essay writing in class, they should be allowed to use the writing process in their class writing tests and for standardized tests at the state or provincial level. If students are taught in a nonconstructivist manner—for example, by teachers explaining a topic through lecture then asking students to practice the information—we cannot test those students using constructivist prompts in which students gather their own meaning their own way. Such experiences are not authentic to what the students experienced and the grade earned wouldn't be accurate.

Format Tests for Efficient Grading

If you're not using the double-column answer test mentioned earlier in this chapter, still ask students to record their answers on an answer sheet. If we teach more than one period of a subject, we only need to make one class set when we give a test. When we grade such tests, we have to carry home a much smaller set of papers—the answer sheets only, plus one copy of the test. As we grade, we don't have to scan through the test pages looking for the answers—we just run our eyes down the answer sheet.

There are two big problems with this, of course: 1) students don't have the questions in front of them when we return their papers, and 2) students don't always copy answers to the answer sheet correctly.

To solve the first problem, we photocopy enough copies of the test to give them out to students when we return their papers. It makes no sense to use a color-in-the-bubble test that is easily graded but offers no insight to students. What does a score of 18 out of 25 teach a student who cannot reflect on the test prompts and his or her responses to them? This may require a lot of paper, but find a way to give students copies of the tests when returning the answer sheets. If we can, save a lot of trees and photocopying costs and post the questions on the class Web site after the test has been given. It doesn't matter that next year's students see the questions in advance. They can do that with paper copies, too. What matters is that the questions are good enough to ask.

To solve the second problem, ask the student to pause at the turn-in basket before inserting the test and go over it one more time, connecting the intended answer with what he or she wrote on the answer sheet. For students with learning disabilities, ask them to write the answers on the test booklet, then help them copy responses to the corresponding blanks on the answer sheet. This is not cheating, and it leads to efficient grading.

We can make our multiple-choice, matching, or true/false questions have responses that create a pattern when recorded, too. This makes for easy grading. For example, an answer pattern might be, "dabadabadaba" ("daba" three times in a row). We can quickly see a letter that's out of the sequence. Of course, the moment our test answers reveal a pattern, we have to throw a curve ball so that students don't just record answers according to the pattern. For example, here's a successful answer pattern for true/false responses—TFFTTFFFTT—and here's an unsuccessful pattern—TFTFTFTFTFTF.

Use Smaller Tests Over Time

In order to get an accurate rendering of students' mastery and support the emphasis on formative assessment mentioned earlier, smart teachers give multiple, smaller, and focused tests over the course of the grading period, instead of one large test at the end. They do this for two reasons. First, that one day of testing at the end of the grading period can have a zillion factors negatively impacting students' performances. Testing is already a snapshot-in-time inference of a student's skill development. Why skew our interpretation further by limiting the opportunities and angles with which students can share what they know? If "all our eggs are in one basket," they can be crushed, never to be recovered, by one clumsy act. Grades are too important to students and their families to not diversify the portfolio from which decisions are made.

Perhaps more important, the more curriculum we put on a test, the less reliable that test grade is in providing specific feedback to students and teachers regarding what it is assessing. In one test, students may define vocabulary, make connections, analyze concepts, demonstrate memorized material, apply knowledge to new situations, and sift text for salient ideas—all in multiple content areas, and only if they interpret the directions for each prompt correctly. If we use longer tests that assess more than one skill or content area, it is wise to record more than one grade at the top in order to more accurately reflect the students' achievement and increase the usefulness of the grades.

One last caution: If students are asking us to hurry up and give them the test before they forget the material, are we teaching for long-term learning? Are students learning the material for the sake of the test alone? If so, what can we do to help them see the material's significance beyond the test? What makes students perceive their learning as fragile?

The brain can reach a saturation point where it feels like it has no more room for storage. We talk ourselves into this condition every time we sit and listen to a lecturer drone on for hours. Coherence weakens, neurons are pruned, and ideas get mixed together. If we're nearing our saturation point for material on which we will be held accountable, we get nervous. To allevi-

ate some of this anxiety, it's appropriate to provide frequent opportunities and motivation to process and permanently store the information in our minds. Preparing for a test provides both.

Include Two Special Questions

First, on the tests themselves, ask students, "What did you think would be asked on this test but was not?" and as appropriate, provide the follow-up prompt: "How would you answer that question?" These questions are not necessary, of course, if students have received a copy of the test at the beginning of the unit, but it makes a great question if the students haven't already seen the test.

Second, include a question that at first sounds reasonable, but on closer examination, is impossible to answer. You'll get a good sense of a student's understanding by how he or she responds; that is, by deferring the response as impossible and explaining why, or by attempting to answer by bluffing through the response. Tell students such a question exists and not to be surprised by it, and give students plenty of experience responding to such questions while teaching the unit.

Tier Questions as Warranted

If some students focused on a different number of objectives or a different level of instruction during the unit, offer assessment questions of varying sophistication in each section of the test, and ask students to answer only the questions identified for their level. An alternative is to design one large test with all the questions, then circle the particular questions you want individual students to answer.

Is it okay for early readiness students to attempt the more complex questions? Sure. Ask them to answer their own level questions first, however, before attempting the more complex ones. Of course, successful responses to the more complex questions would require an altered scoring approach. They would also indicate a need to change your instruction.

Differentiated or not, it's wise to record the learning outcomes or standards the test is assessing at the top of each. If we level or tier tests in any way, we reflect that in the amount or wording of standards recorded. Recording the standards at the top of our tests keeps us and our students focused on the learning, not just the number of problems correct and incorrect. It also helps us keep track of students' achievement.

We may also want to consider how we sequence test items. Some of us prefer to start with relatively easy questions early in the testing sequence.

This gets students warmed up for the more complex responses required later, we reason. Others prefer to mix up the challenge index by placing test items requiring complex responses early in the test or by spacing them evenly throughout, rather than lumping them all at the end. This helps keep a proper attitude toward the test items, we think; and students don't get overwhelmed or intimidated by the complexity of the last questions, nor are their minds tired just when the going gets tough. We tell students in classes in which we do this to move beyond a test item that has them initially stumped, because something in later items might strike them as helpful in solving it.

Either approach can be successful, but test fatigue and item intimidation can be formidable. From my own experience, it is preferable in most tests to spread the challenge throughout, not clump it at the end. Note that assessment expert Rick Stiggins disagrees with this practice, however, calling instead for arranging items from easy to hard (Stiggins et al. 2004, p. 151).

Closing Thoughts

Increasing or decreasing rigor (or preferably, "vigor") in testing does not mean changing the number of tests or test items. It refers to increasing or decreasing the complexity or challenge of the required responses—tiering. We've discussed multiple ways to tier assessments, but three important factors must guide test design.

First, we make sure the question formats don't impede students' successful demonstration of mastery. Anything that might thwart a student in his or her response, such as confusing negatives, tiring matching arrangements, and prompts or answer choices that force students to play "guess what's on the teacher's mind," is immediately discarded.

Second, we level tests and quizzes for students' readiness. All differentiated instruction begins with a fair and developmentally appropriate curriculum, which includes assessment. Students won't learn any faster or better by being pushed to respond to assessments that are not geared for their developmental level. If they're ready for that advanced "pushing" by the teacher, that's great—it's developmentally appropriate.

Just as we might do when forcing a square block into a round hole, something will have to be removed from the student if the assessment format doesn't fit the child's needs. The student's mastery and motivation are diminished by forcing the fit. Instead of doing better, the student may do worse in the long run, failing the test and believing he or she is not capable. We will spend more time and energy overcoming that negative situation than we would spending time designing appropriately leveled assessments.

Does this mean we might increase our record-keeping, such as keeping one gradebook with grades, but another record book of standards and bench-

marks achieved? Yes, though some electronic gradebook programs can do this for you. Does it mean we might need to re-examine our district's report card format or provide a supplemental report card that more accurately represents that student's achievement? Yes.

Finally, we need to get feedback to students in a timely manner. That means we design our tests and quizzes to be graded efficiently, and we make sure students get copies of the tests and quizzes with their answer sheets so they can learn from their mistakes. Some of our tests and quizzes will be in constructed response formats that are impossible to grade quickly, however, especially if we teach more than one hundred students. Quick feedback is still important though, so we try to make tests and quizzes short, such as one-page writings, five sample problems, and oral explanations, so that students get the feedback they need. Feedback is not only information that is used by students, but it's also motivational.