



*a podcast about
how we learn,
how we teach,
and how they overlap*

Episode 38: How to Create Effective Tests

Adam: Hi, I'm Adam Sanford. I'm an academic life coach and professor in Los Angeles.

Dinur: And I'm Dinur Blum. I'm a college professor in Los Angeles.

And this is Learning Made Easier, a podcast where we discuss how we learn, how we teach, and how they overlap.

Adam: Welcome back to Learning Made Easier. This is Episode 38 - How to Create Effective Tests.

Now this is the first in a six-part series for both students and teachers about all things test-related - apart from standardized tests. Those are a whole different animal. But in this series we're going to talk about test creation, question creation, creating and using rubrics, and student strategies for taking tests.

Dinur: This episode is geared towards teachers more than students, though, students who are interested in learning the logic that we use and the reasons we have for writing certain test questions - you're more than welcome to listen in.

Adam: So according to Kansas State University, the type of test you're writing matters, because different types of test questions check different kinds of knowledge. We're going to link to their comprehensive guide in the show notes, and we strongly encourage you to download it.

Now, there are several things to think about, and the first one is basically: What are you testing with this test?

Dinur: And there are two main types of tests: subjective and objective.

Subjective tests are based on how well the student can explain what they know. They should be relatively low-stakes. The students' writing skills should not be the most important thing, because they're taking it timed. They don't get time to proofread as much as they would if this was a take home essay. What you'll receive as essentially a first draft, they're throwing out their ideas. They're often more subjective in the way they're graded, although you can reduce that subjectiveness by using a rubric, about which we're going to talk about in a future episode.

Objective tests are appropriate for impartial or fair evaluation of basic knowledge. These are more likely to use an objective standard for grading, although some questions may not meet that standard. In this case, you'll want to use a key to grade them, rather than a rubric.

Adam: Now, when we are looking at “what are we testing?”, we should probably use Bloom's Taxonomy, which was a tool that was developed years and years ago basically to say “these are the different levels of learning, the different levels of knowledge that we want to see.” And Bloom's Taxonomy is used in the K-12 system. It's used in colleges. So it's not like something that only applies to college students and college tests. But knowing Bloom's Taxonomy can often give you guidelines about: what am I testing, how am I testing it, what's a good way to test it, and how do I know if my question isn't actually testing what I think I'm testing for?

So both subjective and objective tests, they're very useful to measure a student's ability to apply principles, measure their knowledge, demonstrate their critical thinking, and show that they've got problem-solving skills. But different kinds of questions measure different kinds of skills and knowledge. So in this episode, again, Bloom's Taxonomy of knowledge is a great guideline, and it has six steps or levels.

Dinur: And the first, most basic level of Bloom's Taxonomy is called “knowledge.” At this level, we're checking a student's recognition of information and their recall, their memory of that information. At the knowledge level we're checking whether they can remember facts, dates, events, people, terms, theories and so on and whether they can define these ideas and these concepts.

Adam: Now, the second level is called “comprehension” or “understanding.” And in this level we are checking a student's ability to understand what they've learned. So can they explain it? Can they interpret it? Can they restate it in new words? These are common ways to check the second level of Bloom's Taxonomy.

Dinur: The third level is “application.” In this level we ask can the student take what they've learned and use it to explain or understand new information? For example, can they take a specific example of something and identify which group of things it belongs to? Can students use a formula to solve a problem they haven't seen before? Can they use a pattern or a scaffolded idea to create something new?

Adam: At the fourth level, we get to “analysis.” Now in this level, we want the student to show that they can identify how a system is set up by identifying the different parts of the system, or pieces of the system, and explaining how they go together. For example, a student might be asked in a history class or in a civics class to analyze the three branches of the American system of government. Well, then, this would involve defining each branch, and then explaining what powers it has, and explaining what limits other branches have on it, and how they all connect to make this system called “government.”

Dinur: The fifth level is “synthesis.” At this level, students should be able to make their own connections and come up with new perspectives and new patterns from other information. For example, a student who can take two ideas that seem unrelated at first glance, and then show how they relate to one another, is performing synthesis.

Adam: Finally, the sixth level, “evaluation,” asks the student to demonstrate their ability to take some idea and extend it out to its effects. So: “Here's a proposed solution to homelessness. How would it work, and what effects would it have on the city, the community, the homeless community policing practices and so forth?” A question like this demands evaluation.

Dinur: Now, some kinds of test questions are limited to one or two levels of Bloom's Taxonomy. Adam and I will identify which questions test which levels as we go through those questions in the next few episodes.

Adam: So the next step that we're talking about now is, creating tests. And we're not talking about the questions yet, even - we're just talking about the test itself.

So before you create a test, you need answers to the following questions:

What knowledge and skills is this test going to be testing?

Second, how many questions are you going to put in the test?

And third, how many formats are you going to use in the test? How many different kinds of questions are you willing to put in there?

Dinur: Make sure that you're using language that's appropriate to the test and to the students. And that means you don't use figures of speech, you don't use clichés, but you also don't use elevated vocabulary. You want to be clear and concise. You're testing knowledge, not whether they know a dictionary's worth of words, unless you're literally teaching vocabulary at that level.

You also want to remove sex and gender, racial, and ethnic biases from your exams. Unless you're specifically testing knowledge of these biases, keep them out of your questions to keep your tests as neutral and as fair towards their students.

Adam: Make sure that you have actors in your questions that aren't all white males. Make sure that women, non whites and other minorities also appear in the questions. You can also make questions gender neutral by using "they" instead of "he and she," and using gender neutral names like Sam, Chris and Pat, which could be either boys or girls.

You also want to make sure that your test questions are focused on skills and knowledge that were brought up and taught in the lessons used for creating the test. If you give a test that covers stuff that was not covered in class, that's not fair to the students.

The last thing in this area is that you've got to figure out about how much time it's going to take to take the test. So here's good rules of thumb, and these again are directly from the Kansas State handout, which we really recommend that you download. When Dinur and I read it - we're still not finished reading all of it, but when we read it, we're like, "Oh, this is like six different episodes right here. This is great!" - and so we wanted to tell you about it.

So the Kansas State handout recommends this for determining how much time it takes for a student to take a test. For each true false item, allow 30 seconds.

For each multiple choice item, allow 60 seconds or one minute.

For each short answer item, allow 120 seconds, or two minutes.

For an essay question, allow 10 to 15 minutes.

And allow students five to 10 minutes to review their work afterwards.

So you could put this together in a spreadsheet and say, "Okay, I'm going to allow 10 true-false items and 20 multiple-choicers. All right, well, 20 multiple choicers is 20 minutes; and 10 true-falsers, that's five minutes."

Now we've got 25 minutes. If you know that you've only got 45 minutes for the students to take the whole test, then you can only, maybe, have one more essay question, maybe two, tops, because you've got to allow them some time to review what they've done before they turn it in.

Dinur: Now, I tend to do a little bit of a different structuring on my tests. I admittedly will do - because my tests tend to be multiple-choice and true-false, I usually offer students about a minute and a half per question. So if I write a 50-question exam, I give them 75 minutes to take the exam, because that's a minute and a half per question. If I write a 40-question exam than they have an hour, because at that point I'm testing how quickly they can work with information. My tests tend to not focus on memorization a lot. Instead, I want them to be able to apply ideas, and I want to see that they're able to do that in a fairly quick and timely fashion.

Adam: Mm-hmm. And for me, I will take that and say, "all right, for every step they have to take in addition to memorization, I add 30 seconds." So if I've got a multiple-choice question, but it's an analysis level question, then that's not, that's not a one minute question. That's going to be a minute and a half, because they have to do some analysis. So that's an extra step.

Dinur: Now the next part is, you want to make sure your test is fair, and Adam alluded to this a few minutes ago where you said each question should reflect the objectives of the lesson it's related to.

If you're testing stuff that you haven't gone over in class, you're being unfair towards your students, because you've gone over one set of material and you're testing them on something that they have not had a chance to get familiar with.

Each question should present a single task that is clear to the students.

No question should help students answer other questions on the test - so, you don't want to give the answer to number 5 in the prompt for number 12 - and points should be assigned to each question before students take the test.

So for me, I just make each of my questions worth one point, because I don't know what will be easier or tougher for students, and so I don't think it's fair to weight it that well. But if I was offering an essay component on my test, I would weight that more, because that shows me a deeper level of understanding and a deeper level of thinking about the material.

Adam: Now, students often use rules of thumb to help them guess which answers are correct, and badly written tests allow them to do exactly that. So here's how to stop them from doing that.

First, a lot of students assume that the longest answer is the right answer. And sometimes it does have to be, but try really hard to make sure it's only the right answer about a quarter of the time, assuming that you have four options for each question.

Second, make sure that each option number, or each option A, B, C, or D, is used the same number of times as others. There was a slam poem that I saw a couple of years ago. It was teenagers saying, "if you can't

figure out what the right answer is, just fill in C!" And the thing is, because we don't want to give the students the right answer on the first or last option, we tend to default to lots and lots of questions that have "B" or "C" as their options, and C even more than B, because we want it to be wrong answer, wrong answer, right answer. The problem is a lot of students, they'll just fill in C. So make sure that you use each option number about the same number of times as the others.

Third, try to make sure the questions that have "always" or "never" in the correct answer are used about a quarter of the time because students will look at that and say, "any question that's got always or never is always wrong, so I'm not going to answer that." Make sure that about a quarter of the time that is actually the right answer.

Fourth, make sure that answers that offer opposites rarely have either option as a correct answer, because a lot of students will say, "Well, Option A says that this thing is true. Option B says, this thing is false. Option C says, this thing is true when these conditions are there, and Option D says, this thing is true when those conditions are there." One of your correct answers is actually C or D. So make sure that questions that offer opposites, either this or its opposite, rarely have either of those as the correct answer.

Dinur: Number five, try to use five options when possible, to discourage guessing - or deduct an extra point for an incorrect answer, as opposed to not answered, to discourage guessing.

Number six, use scientific-sounding language in the wrong answers as well as in the right answers.

Number seven, about a quarter of the time, make the simple, obvious answer the right answer.

Number eight, use words that are related to the question incorrectly in some answers, so students can't use the words as a guessing basis.

Adam: Now it's only fair to prepare the students for the test. Dropping them in the deep end without giving them any preparation is not fair, but that doesn't mean that you have to have a study session before the test. They should be studying on their own.

But what you can do to help students prepare for the test: give them quizzes that have questions that are formatted the same way as the test, so that they get practice with how you ask questions.

My intro quizzes, that I give the students every day when they come in, that I check their reading, those are pulled directly out of the test bank. Those are actually questions that are in the test bank. Sometimes they'll encounter the same question again. And now, they know what my questions look like. They know what to look for so they can practice how you ask questions.

Another thing is you've got to tell them ahead of time what content will be covered on the test. No, you don't have to give them a study guide, but you do have to guide their study. You could say, "Here's a list of 22 items, and you need to know them all." Or it could be general, "Everything in chapters eight and nine, you need to cover that - because the test is going to be on that stuff."

Also, tell them on the test how many points each question or section is worth, so they can budget their time. If they don't know that the essay is worth 10 times any of the multiple choice questions, they may not spend as much time as they need to on the essay, because they're stressing out about two multiple-choicers that are just driving them batty. They could have gotten at least twice that just for even trying on the essay.

And also, don't penalize for handwriting if they have to take this offline. Suggest that they write in block print if they can, but get used to having to read a lot of different handwriting styles. Some students write clearly, some students, well, don't.

Dinur: And now we come to teachers checking your own work. How do you know when you've written a bad test question and how do you handle it?

And this comes straight from my experience. I sometimes write test questions the way that I talk, and that means that I've written test questions that make perfect sense to me, because I can hear it in my voice and what where I'm putting an inflection. Okay, but that's not necessarily going to make sense to the students taking the exam. This is a bit of a problem, as I don't have to take my own test. And if students can't understand the question, then chances are they're going to answer it incorrectly, because I wasn't communicating what I was trying to test them on in the question very effectively.

So what I do is I look at my question breakdowns, and if over half of my students missed one question, I'll throw that out. I'll award everyone that extra point. You might wonder why.

If so many students miss a question, it means one of a few things. One, okay, maybe my students didn't study that material. Two, or it could mean I didn't cover the material in enough depth in class, nor did I explain a concept from a reading. Or, number three, I just wrote a bad test question. Maybe I'd gone over the material, but the way I wrote the question didn't get at what I was trying to test. Well, two out of those three options place the responsibility for that question on me as a teacher. So to me, that's not fair to hold against students.

Now, that said, I'll make sure to cover the question or those questions once students are done with the tests, so that that information does get covered. But in the meantime, their grades have not been hurt by the mistake.

Now, I don't like to repeat my mistakes - I like to make new ones. (Adam laughs)

So for future tests, I'll try to rephrase awkwardly phrased questions, and I'll try to write, I'm working on writing scenario-based questions, where my students now have to apply concepts or definitions, and answer questions based on these given scenarios. And that means that students have to analyze what the questions ask them, and analyze information, and be familiar enough with material from class to be able to work with it. This is opposed to what I used to do, of asking students for a strict definition of a term, because it's less useful, I think, to know the strict definition than to be able to work with this idea.

Adam: So another thing you might want to do, and I do this - give your students a way to notify you about bad or confusing test questions. So if they're taking a paper exam, and you're okay with them writing on the exam, then have them put a symbol on the exam. Next to questions that don't make sense, like just have them put a square around the, uh, around the question number to indicate "this question makes no sense to me. I couldn't figure it out." And then afterwards, go through, you know, collect all the test papers, and then go through the test papers and say, "all right, you know, out of the 35 exams, we have 18 of them with question 12 has a square around it. Something's wrong there, I probably screwed up."

If they can't write on the exam then maybe give them, uh, you know, let them - I would actually walk around the class and put, like, a Post-It note on everybody's desk after they've started taking the exam so that they can write down, on the Post-It note, "this question made no sense to me. Number 12 made no sense," right?

Because my exams are online, I make a point of asking students to email me a screenshot of a test question if it doesn't make sense, and explain why it's confusing. And if I get a question like that, I will give the points to all students for that question, and then I'll revise them out of the test bank for the next time I use it.

So as an example, I recently realized that leaving out the word "only" completely changed the meaning of a question in ways I didn't intend it to be changed. But I owned the mistake, and I gave everybody credit for it. Another question - and this happens - there was a mistake in the key! And again, since it was my mistake, I owned it and I gave credit and I said, "By the way, the right answer for this is..."

This is also a really good way to model effective mistake-making, and how to handle mistakes, for your students, so, win-win.

Dinur: Yep. This actually just happened to me with a test I gave recently, where I thought I'd written one of my answer choices clearly, and I didn't. And so two of my answer choices repeated themselves, and a few of my students who took the test early pointed it out to me, so I was able to revise before most of my students had taken it and I immediately gave those students the points back for that question. And I explained, "here's what I meant for this answer choice to be, and here's why this was the correct answer."

But again, that's an error on me. But I'd rather be open and honest, and I know Adam is the same way - be open and transparent and say, "okay, this was a mistake. We owned it. Let's move on." There was no harm done because the points have been awarded.

Adam: Right.

Dinur: So that's what we have for you in Episode 38. If you're finding this podcast helpful, please share it with your friends! We're always hoping to get new subscribers, so we can help more people. You can find us on Apple Podcasts, Spotify, and Android. We're hosted on Blubrry.com. Also, we'd appreciate it if you wrote a review of this podcast on Apple Podcasts.

Adam: Be sure to join us next week for Episode 39: how to write effective multiple-choice, true-false, and matching questions.

Dinur: You've been listening to Learning Made Easier, a podcast about how we learn, how we teach, and how they overlap.

Adam: We want to say thank you to all of our supporters on Patreon, who make this podcast possible.

Dinur: If you want to support us, please go to www.patreon.com/learningmadeeasier.

Adam: We look forward to seeing you next week!