Essay Items

Hope You Can Write

Difficulty Index ☺☺☺☺☺ (really pretty easy)

One of the things that we will stress throughout this part of Tests & Measurement for People Who Hate Tests & Measurement is that the type of item you use as an assessment tool is closely tied to the level of information (or level of thinking skills) you want your test taker to show that he or she does (or does not) have. In Chapter 6, we emphasized how short answer items are best used when trying to assess memorization of factual information. In this chapter on essay items, you’ll find other tools that you can use to assess other types of outcomes.

Here, we’ll focus on essay items—those items where the test taker is expected to write a coherent and informative response to a question. Forget about that Friday spelling test or even the SAT—essay tests are the real thing when it comes to seeing how well test takers integrate ideas and how well they can express them in written form.

When We Use ‘Em and What They Look Like

Essay questions allow for perhaps the most unrestricted type of written assessment item that we will cover in Tests & Measurement for People Who Hate Tests & Measurement. What you want to know is how well the test taker can organize information and express his or her ideas in writing. That’s why the really, really big exams in one’s academic career are usually of the essay type—these types of items just tap more higher-level and complex skills.
Essay items are the item of choice if you want an unrestricted response and want to access higher-order thinking, such as the relationship between ideas and the pros and cons of a particular argument.

Essay questions come in two basic flavors: open-ended (also called unrestricted or extended) questions and closed-ended (also called restricted) questions.

An open-ended (or unrestricted response) essay question is one where there are no restrictions on the response, including the amount of time allowed to finish, the number of pages written, or material included. Now, it is a bit impractical to allow test takers to have 25 hours to answer one essay question or to write hundreds of pages about anything they like. So, of course there are practical limits. It's just that the limits do not define the scope of the response. For example, here's an open-ended essay question.

1. Discuss the various theories of human development that have been talked about this semester. Among other things, be sure to compare and contrast the basic assumptions of the theories, the method used to study development, and the criticisms of each. You are free to take as much time as you need. Write as many pages as you like, and organize your answer as you see fit.

That is a nice, open-ended question where the response is unrestricted.

Now, take a look at this closed-ended question where the response is restricted.

1. Compare and contrast two of the basic theories of human development that have been discussed this semester. Include a brief summary of each theory, the method used to study development, and the criticisms of each perspective. Limit your response to five written pages, and do not write for longer than 2 hours.

These two types of questions reflect different types of experiences. The first question, which is much less restrictive, gives the test taker a lot more flexibility (among other things) and allows for a more creative approach. The more restricted closed-ended question places definite limits on the content as well as the format.
You might think that almost everyone would like to have as much flexibility as possible, but that’s just not the case—many people like a very well-structured and clearly defined task assigned to them.

**HOW TO WRITE ESSAY ITEMS: THE GUIDELINES**

Here are just a few guidelines that will be helpful when it comes time to write an essay question.

1. *Allow adequate time to answer the question.* By their very design, essay questions can take a considerable amount of time to answer. Earlier in this chapter, we contrasted closed- and open-ended questions, where the nature of the response is restricted or not. Regardless of whether an essay question is closed or open ended (remember, we have to be practical), you need to tell the test taker how much time he or she has to complete the question. And how much time should that be? Keep in mind that essay questions require test takers to think about, and then write, the response. One strategy is to encourage test takers when they are practicing to plan their response by spending 30%-40% of their time outlining or “sketching” their response, 40%-50% of their time writing the response, and then the last 10%-30% rereading what they have written and making any necessary changes.

**Using Outlining Tools**

These days, it is not unusual for a test taker to complete an essay using a word processor and not even manually write anything. Test takers in such a situation may also be able to use the outlining feature that many word processors provide. The 30% or 40% of time spent thinking about the response can be used to outline the important major and minor points. Then, the outlining features—such as expanding and contracting headings, moving headings up or back in the outline, and collapsing and expanding sections—can be used to fine tune the response before the actual writing begins. Some essay writers write out their thoughts as an outline and, when finished, create the necessary transitions between thoughts and are done! Figure 7.1 shows you what such an outline might look like.
2. Be sure the question is complete and clear. This one sounds simple and it may indeed be, but sometimes essay questions are not very clear in their presentation. Want to know why? Because it’s not clear what the person writing the question wants to know. For example, here’s an unclear essay question

1. Discuss the impact of the Civil War on the economy of the postwar South.

It’s not like this is that poorly designed a question, but it sure does not reflect a clear notion of what was learned or what is being assessed. This is the kind of a topic that some historian could write seven volumes about! Look how much more clear the following question is.

1. Discuss the impact of the Civil War on the economy of the postwar South, taking into account the following factors: reduction in the work force, international considerations, and the changing role of agriculture.
Chapter 7  ❖ Essay Items: Hope You Can Write

This second example just provides additional direction, which may be exactly what the test taker needs to answer the question more completely.

3. Essay questions should be used only to evaluate higher-order outcomes, such as when comparisons, evaluations, analyses, and interpretations are required. Want to know what $64^3$ is? (262,144)—the infant mortality rate of the United States in 2001? (6.9 per 1,000 live births)—what the French called tomatoes? (pomme d'amour for apple of love). If so, an essay question is not what you are looking for—you want the kind of item that tackles lower-level thinking skills such as knowledge or memorization. For the higher-order outcomes, such as evaluation and synthesis, you want the essay format and questions such as the following:

1. Identify the principle factors that contribute to infant mortality in the United States and discuss and evaluate the success of efforts that have been taken to close the gap between white and nonwhite populations.

Such a question requires a knowledge of rates of infant mortality but goes far beyond the statement of just facts. In fact, here’s this chapter’s cheat sheet about what kinds of words signal higher-order skills (and these words usually show up in essay items).

<table>
<thead>
<tr>
<th>Analyze</th>
<th>Generalize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply</td>
<td>Infer</td>
</tr>
<tr>
<td>Arrange</td>
<td>Integrate</td>
</tr>
<tr>
<td>Classify</td>
<td>Justify</td>
</tr>
<tr>
<td>Contrast</td>
<td>Organize</td>
</tr>
<tr>
<td>Compare</td>
<td>Persuade</td>
</tr>
<tr>
<td>Create</td>
<td>Predict</td>
</tr>
<tr>
<td>Evaluate</td>
<td>Summarize</td>
</tr>
<tr>
<td>Explain</td>
<td>Synthesize</td>
</tr>
</tbody>
</table>

4. Have all test takers answer the same questions. This just reduces the burden placed on the developer of the test, but more important, makes the test easier to score—it's just more practical. Why?
Because essay questions should have one correct answer, and (as you will learn shortly) there should be a model answer against which test takers’ responses should be compared. Allowing for multiple questions requires the test scores to have multiple standards against which to compare them—too much work! Instead, rotate questions from year to year or test to test so there is a nice collection of questions and model answers.

**THE GOOD AND THE BAD**

Essay items have their advantages and disadvantages—let’s review them. And, you can find a summary of these in Table 7.1.

<table>
<thead>
<tr>
<th>TABLE 7.1 The Advantages and Disadvantages of Essay Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages of Essay Items</strong></td>
</tr>
<tr>
<td>• They help find out how ideas are related to one another.</td>
</tr>
<tr>
<td>• They increase security.</td>
</tr>
<tr>
<td>• They provide increased flexibility in item design.</td>
</tr>
<tr>
<td>• They are relatively easy to construct.</td>
</tr>
<tr>
<td><strong>Disadvantages of Essay Items</strong></td>
</tr>
<tr>
<td>• They emphasize writing.</td>
</tr>
<tr>
<td>• They are difficult to write.</td>
</tr>
<tr>
<td>• They provide an inadequate sampling of subject matter.</td>
</tr>
<tr>
<td>• They are hard to score.</td>
</tr>
<tr>
<td>• They emphasize writing skills over content.</td>
</tr>
</tbody>
</table>

**Why Essay Items Are Good**

1. It’s doubtful that you can come up with a better way of finding out not only what test takers know, but also how they relate ideas to one another. You’ve defined your learning goals, and the essay question is your choice of assessment tool because you are interested in finding out how well a test taker understands ideas and can relate ideas to one another. Essay questions most definitely tap into how well test takers can organize and integrate information. And essay questions also provide opportunities to demonstrate creativity.
Chapter 7 • Essay Items: Hope You Can Write

THINGS TO REMEMBER: WHY ESSAYS?

Among the many reasons why essays are particularly useful is that they are very flexible in both form (the size and complexity) and purpose (relate simple ideas or elaborate a complex argument).

2. The use of essay questions increases security because it is very difficult to plagiarize during an essay item examination. And, along the same lines, it is almost impossible for test takers to effectively guess the correct answer—so guessing is removed as a legitimate concern.

3. The essay item format has unparalleled flexibility. Take the time necessary to create a really good question—one that is exciting to consider and one that gives the test taker a fair chance at success.

4. Essay questions can be relatively easy to construct. If you know your material well (and we would sure assume such), you can put together four essay questions in less than an hour that can effectively tap higher-order learning. Sure beats creating 100 multiple-choice items at 10 minutes each (yep, about 15 hours).

Why Essay Items Are Not So Good

1. They emphasize writing. Well, no kidding—that’s what they are supposed to do. But what they don’t do well at all is tapping the test taker who is knowledgeable about ideas and their relationship to one another, but just cannot express it in words. For these kinds of test takers, evaluating them fairly and accurately is always a challenge.

2. They can be tough to write. But didn’t we just say above that they were easy to write? Well, we said they can be. They can be a bear as well! The test designer has to invest a great deal of time in creating each essay question and making sure that it taps the objectives that are to be tested, but does so in a way to encourage the expression of ideas that you can’t find in a true/false or matching test.
3. **Precision in sampling counts.** Because essay questions take a good deal of time to create and even more time to complete as part of a test, it's tough for the test to adequately sample the entire universe of what the test taker might have learned. This should always be a consideration in the overall evaluation of knowledge and understanding of any topic.

4. **Essay questions are not easy to score.** Think about it—in a class of 25 students, each of whom completes five essay questions—that's more than 120 questions that have to be graded. No teacher will begrudge grading them (well, maybe a few), but all will have some serious problems remaining neutral, staying on task, and being consistent.

5. **Writing can become more important than content (see #1 above).** Because essays are written, some students can bluff their way through the answer by virtue of their excellent writing ability. Just about the only way to counter this is through the use of model answers, as discussed in the next section.

---

### HOW TO SCORE ESSAY ITEMS

Essay items are absolutely terrific and almost indispensable for sampling higher-order thinking. I've mentioned that several times throughout this chapter. But they are a bear to score—time consuming and very demanding of all the scorer's attention.

With that in mind, here are a few tips that might make the scoring process more efficient and result in a fairer assessment.

1. **Scorers should provide plenty of time to score an essay item.** Each of the items has to be read and then scored, and often the scorer will read the items more than once—the first time for a general overview of the content, and the second time for a more detailed analysis, including an assessment of content (again) and writing skills (such as grammar, transitions, and sentence usage).
There’s one thing about scoring essay items that you have to be especially careful of, and that’s allowing yourself to be swayed by the very first question you evaluate. If the handwriting is difficult to read, if you inadvertently know the test taker (which may be very easy given the handwriting), and so on—you may be biased one way or the other. Biased about what? All the grading that follows. A very good, or a very poor, response can create expectations (which may or may not be appropriate), and it’s not easy to shed those as you grade. Unless you are able to approach some (acceptable) level of objectivity, then you should seriously consider creating a model you can follow point by point, argument by argument.

2. *Take your time.* This is related to #1 above, but it is critical to pace yourself when scoring essay items. If you have 50 students, each writing four essay responses, it would be impossible to grade these 200 (yes, 200) responses in one sitting. Fatigue is a huge threat to the integrity of the assessment process, so grade essay exams in batches. Perhaps one question at a time, with 30 minutes or something else in between (like a walk around the block, or a stretch, or get yourself a cookie as a reward for your good work!).

3. *Use a model correct answer to have a basis for comparison.* Having a model greatly increases the likelihood that you will evaluate each answer fairly and have as objective a standard as is possible because you can compare what is there (the test taker’s response) to what should be there (your model response). Answering your own questions is not fun, especially when you know you have to do a terrific job, but this is about the only method you can use to avoid a totally subjective approach to scoring these kinds of questions. When’s the best time to create these model answers? When you’re initially writing the question, of course!

You can also develop your own system of scoring. Susan M. Brookhart, a grading expert from Duquesne University, suggests that essay questions should be evaluated on three criteria:
5. If possible, grade the responses without knowing the test taker’s identity. Because a subjective element can enter into the grading of essay questions, not knowing who the test taker is (and avoiding that possible bias) can be a great help. There are a bunch of ways to do this (using code numbers keyed to names rather than just names is one example), but another way is to have the test taker put his or her name on the back (and blank) side of the first page of the responses. The grader won’t see it and any conflict can be avoided.

4. **Score each question across all test takers.** Given that you have a model answer, go through each test focusing the same question each time. For example, score the answer to Question 1 across all test takers, and then go back and score Question 2 across all test takers. This allows you to make absolute judgments in comparison to your model answer, but also to make relative judgments (if necessary) within any one item.

---

**Essay Items and Unreliability**

You remember from Chapter 2 that our discussion about reliability focused in part on the reduction of error variance. The more error variance that is removed from the entire testing situation, the higher the reliability.

OK—given that thought, there are a lot of sources of error variance when it comes to essay items, and perhaps the largest source is differences in grading. By its very nature, an essay item has so many indefinite things associated with it (objectivity of the grader, student’s writing skills, etc.) that it is even more important to control what you can. So, do try and ensure anonymity, use a model for scoring, and standardize conditions whenever possible. You’ll have a more reliable test, and the test takers will be treated more fairly.

---

• Thesis and organization
• Content knowledge
• Writing style and mechanics

Each of these elements, for example, could be scored separately. A model too difficult for you to do? Then at least have the important points listed on a separate sheet of paper against which you can compare test takers’ responses.
SUMMARY

Essay questions are terrific to get at those more sophisticated thinking skills that we have to assess to find out if an individual understands more complex ideas and how these ideas relate to one another. But although they might be relatively easy to create (they are short, after all, and only a few are needed), they can underrepresent content, they are tough to score, and there are lots of reasons why they could be considered unreliable. So, use 'em only when they need to be used!

TIME TO PRACTICE

1. In your area of interest, write one stunningly terrific essay question. Then, exchange it with a classmate and evaluate it according to the guidelines that were presented in this chapter.

2. Explain why essay questions are more useful for tapping higher-order thinking skills than just straight memorization.

3. Name two advantages of essay items and provide an explanation as to why they are advantages.

4. Write three essay items that violate at least one of the guidelines we identified in this chapter, and indicate what's wrong.

5. Name two disadvantages of essay items and discuss how you might compensate for them.

ANSWERS TO PRACTICE QUESTIONS

1. Here's my question:

   Discuss the origins of the testing movement in the United States and be sure to identify how the social and political events of the times helped encourage the growth of the testing industry (45 minutes).

2. Unlike test items such as short answer or completion (or many other types you have yet to learn about), essay questions provide the flexibility to explore ideas and not just require memorization or the repetition of facts.

3. First, good performance on an essay question demands an understanding of ideas and how they relate to one another. That's very important in topics where there is a premium on relationships between ideas and concepts rather than rote memorization of those concepts. Second, they are very flexible in that they can be adjusted to meet the needs of the assessment setting. One set of essays might be used to look at a description of a particular historical event, whereas a more
ambitious one (for the test taker, that is) might be used to look at how certain unhistorical events had an impact on others and what the consequences of each event was for future policy decisions.

4. Here are mine:
   a. What are some of the important things that happened in American history? (unclear and incomplete)
   b. Name the planets in order of their distance from the sun, with the closest first. (not an essay question)
   c. When she didn’t establish a foundation for the foundation, she made a terrible mistake that angered her mother. Relate that to how she furthered her career. (Uh—very unclear and include).

5. My two are as follows:
   a. There’s no getting around the fact that the writing component to essay questions can be very important. But one way to compensate for that is to make sure the test takers are capable of expressing themselves in writing before administering such an exam. Or, provide adequate outlining assistance and extensive practice organizing and writing such a response. These are tough to get by, but perhaps worth it.
   b. Essay questions are no fun, tedious, and very draining to score—especially if you have a lot of them to do. One way to help with this task is to use a model answer (as we discussed earlier in this chapter) and to work in small chunks of time so that you remain refreshed.

WANT TO KNOW MORE?

Further Readings


  Essay questions can be used in many different settings as exemplified here, where these researchers examined gender differences in the diagnosis and management of neck pain. Modified essay questions were used to record suggestions about management of neck pain.


  The title says it all. Students in the top third on the essay portion of a test and the bottom third on the multiple-choice portion of the
examination were compared with students with the opposite pattern (top third on the multiple-choice questions and bottom third on the essay questions). Part of the findings? Students who were strong in the essay format and weak in the multiple-choice format were as successful in their college courses as students with the opposite pattern. Guess if you know it, you know it.

And on the Internet

- The Advanced Placement people at the University of Georgia can show you some real live essay questions from previous years dating back to the early 1960s at http://apbio.biosci.uga.edu/exam/Essays/. This is a perfect place to see what essay questions look like in the area of biology at some of the most advanced levels.
