Item Writing: Tips for Authoring Test Questions

📅 April 12, 2024

👤 [Nathan Thompson, PhD](#)

Item writing (aka item authoring) is a science as well as an art, and if you have done it, you know just how challenging it can be! You are experts at what you do, and you want to make sure that your examinees are too. But it's hard to write questions that are clear, reliable, unbiased, and differentiate on the thing you are trying to assess. Here are some tips.

What is Item Writing / Item Authoring ?

[Item authoring](#) is the process of creating test questions. You have certainly seen "bad" test questions in your life, and know firsthand just how frustrating and confusing that can be. Fortunately, there is a lot of research in the field of [psychometrics](#) on how to write good questions, and also how to have other experts review them to ensure quality. It is best practice to make items [go through a workflow](#), so that the [test development process](#) is similar to the software development process.

...ause items are the building blocks of tests, it is likely that the test items within your ...ts are the greatest threat to its overall validity and reliability. Here are some

important tips in item authoring. Want deeper guidance? Check out our Item Writing Guide.

Anatomy of an Item

First, let's talk a little bit about the parts of a test question. The diagram on the right shows a reading passage with two questions on it. Here are some of the terms used:

- Asset/Stimulus: This is a reading passage here, but could also be an audio, video, table, PDF, or other resource
- Item: An overall test question, usually called an "item" rather than a "question" because sometimes they might be statements.
- Stem: The part of the item that presents the situation or poses a question.
- Options: All of the choices to answer.
- Key: The correct answer.
- Distractors: The incorrect answers.
-

Item writing tips: The Stem

To find out whether your test items are your allies or your enemies, read through your test and identify the items that contain the most prevalent item construction flaws. The first three of the most prevalent construction flaws are located in the item stem (i.e. question). **Look to see if your item stems contain…**

1) BIAS

Nowadays, we tend to think of bias as relating to culture or religion, but there are many more subtle types of biases that oftentimes sneak into your tests. Consider the following questions to determine the extent of bias in your tests:

- Are there are acronyms in your test that are not considered industry standard?
- Are you testing on policies and procedures that may vary from one location to another?
- Are you using vocabulary that is more recognizable to a female examinee than a male?
- Are you referencing objects that are not familiar to examinees from a newer or older generation?

2) NOT

We've all taken tests which ask a negatively worded question. These test items are often the product of item authoring by newbies, but they are devastating to the validity and reliability of your tests—particularly fast test-takers or individuals with lower reading skills. If the examinee misses that one single word, they will get the question wrong even if they actually know the material. This test item ends up penalizing the wrong examinees!

3) EXCESS VERBIAGE

Long stems can be effective and essential in many situations, but they are also more prone to two specific item construction flaws. If the stem is unnecessarily long, it can contribute to examinee fatigue. Because each item requires more energy to read and understand, examinees tire sooner and may begin to perform more poorly later on in the test—regardless of their competence level.

Additionally, long stems often include information that can be used to answer other questions in the test. This could lead your test to be an assessment of whose test-taking memory is best (i.e. "Oh yeah, #5 said XYZ, so the answer to #34 is XYZ.") rather than who knows the material.

Item writing tips: distractors / options

Unfortunately, item stems aren't the only offenders. Experienced test writers actually know that the distractors (i.e. options) are actually more difficult to write than the stems themselves. When you review your test items, **look to see if your item distractors contain**...

4) IMPLAUSIBILTY

The purpose of a distractor is to pull less qualified examinees away from the correct answer by other options that look correct. In order for them to "distract" an examinee from the correct answer, they have to be plausible. The closer they are to being correct, the more difficult the exam will be. If the distractors are obviously incorrect, even unqualified examinees won't pick them. Then your exam will not help you discriminate between examinees who know the material and examinees that do not, which is the entire goal.

5) 3-TO-1 SPLITS

You may recall watching Sesame Street as a child. If so, you remember the song "One of these things…" (Either way, enjoy refreshing your memory!)   Looking back, it seems really elementary, but sometimes our test item options are written in such a way that an examinee can play this simple game with your test. Instead of knowing the material, they can look for the option that stands out as different from the others. Consider the following questions to determine if one of your items falls into this category:

- Is the correct answer significantly longer than the distractors?
- Does the correct answer contain more detail than the distractors?
- Is the grammatical structure different for the answer than for the distractors?

6) ALL OF THE ABOVE

There are a couple of problems with having this phrase (or the opposite "None of the above") as an option. For starters, good test takers know that this is—statistically speaking—usually the correct answer. If it's there and the examinee picks it, they have a better than 50% chance of getting the item right—even if they don't know the content. Also, if they are able to identify two options as correct, they can select "All of the above" without knowing whether or not the third option was correct. These sorts of questions also get in the way of good item analysis.   Whether the examinee gets this item right or wrong, it's harder to ascertain what knowledge they have because the correct answer is so broad.

This is helpful, can I learn more?

Want to learn more about item writing? Here's an instructional video from one of our psychometricians. You should also check out this book.

Item authoring is easier with an item banking system

The process of reading through your exams in search of these flaws in the item authoring is time-consuming (and oftentimes depressing), but it is an essential step towards developing an exam that is valid, reliable, and reflects well on your organization as a whole. We also recommend that you look into getting a dedicated item banking platform, designed to help with this process.

Summary Checklist

ue	Recommendation

Key is invalid due to multiple correct answers.	Consider each answer option individually; the key should be fully correct with each distractor being fully incorrect.
Item was written in a hard to comprehend way, examinees were unable to apply their knowledge because of poor wording.	Ensure that the item can be understood after just one read through. If you have to read the stem multiple times, it needs to be rewritten.
Grammar, spelling, or syntax errors direct savvy test takers toward the correct answer (or away from incorrect answers).	Read the stem, followed by each answer option, aloud. Each answer option should fit with the stem.
Information was introduced in the stem text that was not relevant to the question.	After writing each question, evaluate the content of the stem. It should be clear and concise without introducing irrelevant information.
Item emphasizes trivial facts.	Work off of a test blue print to ensure that each of your items map to a relevant construct. If you are using Bloom's taxonomy or a similar approach, items should be from higher order levels.
Numerical answer options overlap.	Carefully evaluate numerical ranges to ensure there is no overlap among options.
Examinees noticed answer was most often A.	Distribute the key evenly among the answer options. This can be avoided with *FastTest's* randomized delivery functionality.
Key was overly specific compared to 'ractors.	Answer options should all be about the same length and contain the same amount of information.

Key was only option to include key word from item stem.	Avoid re-using key words from the stem text in your answer options. If you do use such words, evenly distribute them among all of the answer options so as to not call out individual options.
Rare exception can be argued to invalidate true/false always/never question.	Avoid using "always" or "never" as there can be unanticipated or rare scenarios. Opt for less absolute terms like "most often" or "rarely".
Distractors were not plausible, key was obvious.	Review each answer option and ensure that it has some bearing in reality. Distractors should be plausible.
Idiom or jargon was used; non-native English speakers did not understand.	It is best to avoid figures of speech, keep the stem text and answer options literal to avoid introducing undue discrimination against certain groups.
Key was significantly longer than distractors.	There is a strong tendency to write a key that is very descriptive. Be wary of this and evaluate distractors to ensure that they are approximately the same length.

Nathan Thompson

Nathan Thompson earned his PhD in Psychometrics from the University of Minnesota, with a focus on computerized adaptive testing. His undergraduate degree was from Luther College with a triple major of Mathematics, Psychology, and Latin. He is primarily interested in the use of AI and software automation to augment and replace the work done by psychometricians, which has provided extensive experience in software design and programming. Dr. Thompson has published over 100 journal articles and conference presentations, but his favorite remains https://scholarworks.umass.edu/pare/vol16/iss1/1/ .

in ▶️

Get in touch, and we'll meet to discuss how we can improve your exam development, delivery, and psychometrics!

Request a Consultation

Join our newsletter

First Name

Last Name

Your email address

Sign up!

Solutions

Test Delivery

Item Banking

otive Testing

Remote Proctoring

Proctored Test Centers

Credicycle

Services

Psychometric Consulting

Test Development

Item Writing & Content

Certification Management

Psychometric Software

Training & Support

Company

About ASC

Leadership

Articles & Resources

Client support portal

Privacy Policy

651.383.4311

solutions@assess.com