

Test Question Development

The ideal question accomplishes three goals:

It measures the underlying concept as intended, it doesn't measure other concepts, and it means the same to all respondents.

Test Basics

- **Fairness:** Content mastery should be measured in a way that does not give advantages to irrelevant factors. Test questions should reflect the learning objectives that were clearly communicated to students. Point values should be determined before the test is given.
- **Accuracy:** Never use a test written by someone else without checking it for accuracy. Publisher questions may save time, but they can be poorly written and answers may be incorrect. Always proofread.
- **Objectivity:** This is one of the biggest challenges since thinking that one type of question is automatically objective and another is undoubtedly subjective is a false assumption. Focusing on the most important content and relying on a second pair of eyes can help.
- **Reliability:** The more items a test has, the more reliable it is. In a short test, a few incorrect answers can have significant impact while a few incorrect answers in a long test makes minimal difference in the outcome.
- **Mixed methods:** A test with a mix of types of questions minimizes student weakness with a particular method of testing.
- **Instructions:** Provide clear, concise written instructions. It is often helpful to provide an example of an exemplary answer.
- **Accommodations:** Think ahead about any accommodations you may be required to give for students with learning or physical disabilities. Consider what accommodations are reasonable.

Question Basics

Write for clarity	The best questions don't require lengthy explanations, creative hypotheticals, or dramatic rewordings – they get to the point right away. Write your questions down, re-read them, and take another stab at it, if necessary.
Write for brevity	Try to say the same idea in fewer words without losing clarity. The sooner you get to the point, the sooner the students can understand your test questions and respond.
Get a reviewer	Find someone whose expertise is close to your own so they can gauge how appropriate each question is for your class and student grade level. If they think a question is too easy or too difficult, take it to heart. You can also follow up with them to see if they had any specific ideas on how to make your question(s) better.
Focus on relevance	Focus on the most relevant content, main theories, or the take-home points - assess application of knowledge, not recall of an isolated fact. That will ensure that you measure and grade the concepts that are most important. Moreover, test questions also serve as a reminder for students and challenge their thinking, which solidifies memory.
Ask one thing at a time	Including two topics in one question is referred to as a double-barreled question. It will not only confuse students, but also muddle your data. Be specific in what you want to know and focus on one thing at a time.
Be concise	A general rule of thumb is to try to ask as few questions as possible. This reduces the burden for the participants and will streamline the analysis.

Main Question Types

Multiple Choice questions

- Question/Stem

Be clear	The stem must be clear and unambiguous. Any vague terms (e.g., normally, usually, possibly, may) or absolutes (e.g., always, never) should be avoided.
Be direct	State the stem as a direct question rather than an incomplete statement. The stem should be meaningful by itself. Avoid stems such as: “Which of the following is a true statement?”
Be concise	Eliminate unnecessary information or excessive verbiage in the stem. The stem should not contain irrelevant material. This can decrease the reliability and the validity of the results.
Be precise	The stem should be meaningful by itself and should present a definite problem. A stem that presents a definite problem helps focus on the learning outcome.
Avoid repetition	Include in the stem any words that are repeated in every alternative.
Avoid negative wording	The stem should be negatively stated only when significant learning outcomes require it. Students often have difficulty understanding items with negative phrasing.
Avoid overlap	Keep the specific content of items independent of one another. Savvy test-takers can use information in one question to answer another question, reducing the validity of the test.

- Effective Alternatives

Be consistent	Make alternatives roughly equivalent in length and in a similar format (i.e., phrases, sentences, etc.). Ensure your alternatives are free from clues for the correct answer.
Be plausible	The purpose of distractors is to assess which students have achieved the learning. Only plausible alternatives serve as functional distractors, which are chosen by students who have not achieved the objective and ignored by students who have achieved the objective. There is little difference in difficulty, discrimination, and test score reliability among items containing two, three, and four distractors. As a rule of thumb, include four alternatives to reduce the impact of guessing.
Be brief	Alternatives should be stated clearly and concisely. Items that are excessively wordy assess students’ reading ability rather than their attainment of the learning objective.
Be unpredictable	Randomly distribute the correct answer among the positions throughout the test, e.g., present the alternatives in a logical order (e.g., alphabetical, numerical) to avoid a bias toward certain positions.
Avoid negative wording	Be cautious in using negative terms in the choices. If you use a negative, capitalize the word for emphasis.
Avoid overlap	Alternatives should be mutually exclusive. Alternatives with overlapping content may be considered “trick” items by test-takers, excessive use of which can erode trust and respect for the testing process.

Short answer questions/Essay questions

- Essay or short-answer questions are easy to write, but time-consuming to grade. Use this type of question when you want to assess the ability to organize knowledge, give an opinion, or use original thinking.
- Avoid giving a selection of questions. This reduces reliability. If students answer different questions, they are taking a completely different test.
- Write questions which are comprehensive, rather than focusing on small units of content.
- To test the depth of background knowledge, require supporting evidence for claims and assertions.
- Communicate how you will respond to technical errors such as misspellings, grammar, etc.
- Identify the relative value of each question.

True or False questions

- Due to the high probability of guessing and blurry distinctions between “true” and “false”, true/false questions are not reliable and not recommended.

Resources

[Link to NBME Gold Standard](#)

[Link to NBME Workshop](#)

Burton, Steven J., Sudweeks, Richard R., Merrill, Paul F., and Wood, Bud. How to Prepare Better Multiple Choice Test Items: Guidelines for University Faculty, 1991.

Cheung, Derek and Bucat, Robert. How can we construct good multiple-choice items? Presented at the Science and Technology Education Conference, Hong Kong, June 20-21, 2002.

Haladyna, Thomas M. Developing and validating multiple-choice test items, 2nd edition. Lawrence Erlbaum Associates, 1999.

Haladyna, Thomas M. and Downing, S. M. Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education*, 2(1), 51-78, 1989.

Karhikeyan, S., O'Connor, E. & Hu, W. Barriers and facilitators to writing quality items for medical school assessments – a scoping review. *BMC Med Educ* 19, 123 (2019).

Morrison, Susan and Free, Kathleen. Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education* 40: 17-24, 2001.