Writing Effective Test Questions Series
PART 2: Writing Selected Response Exam Questions

Selected vs. Constructed Response Test Questions

Selected response test questions are those to which there is typically a single correct answer, and comprise fill-in-the-blank, true-false, multiple choice, and/or matching tasks. Nilson (2016) notes that these questions are good for assessing students' ability to remember and understand course concepts and materials, but cannot "measure students' abilities to create, organize, communicate, define problems, or conduct research" (p. 291). Selected response questions are easily scorable using a machine like a Scantron, which makes them seem like a good choice in large-enrollment courses.

Constructed response questions ask learners to *generate* (or construct) an answer. Constructed response items can measure knowledge, comprehension, application, perspective, and/or self-awareness. Examples of constructed response assessment tasks include: listing, defining, providing reasoning, short answer questions, and essay exams. There primary types of constructed response, restricted and extended, are described below.

Constructed Response Type	Sample Prompts	Pros	Cons
Restricted	Provide reasons for… List… Define …	Allows for faster grading	Does not measure higher-level thinking.
Extended	During this unit, we have discussed both the evolution of American literature and the changing political climate of the twentieth century. Analyze these two dimensions of life in America, citing instances where literature and politics may have influenced each other. Describe those influences in specific terms. In planning your response, think about what we learned about prominent novelists, political satirists, and prominent political figures of the last half of the century. (5 points per instance, total = 15 points).	Measure complex, interrelated skills such as synthesis, evaluation, and expression; as well as knowledge mastery and reasoning proficiency.	Difficult to write well. Challenging to score equitably; requires written English proficiency.

This resource will focus on suggestions for designing multiple choice questions (MCQs). For strategies for designing other types of selected response assessment questions, see Nilson (2016).

Strategies for writing multiple choice test questions

Research suggests that while well-designed multiple choice questions (MCQs) can be used to assess multiple dimensions of Bloom's Cognitive Process Domains, most MCQ tools focus on lower-order skills like remembering and understanding (Momsen et al., 2010). However, well-constructed MCQs can be used to assess higher-level thinking such as apply or analyze (Clegg & Chasin, 1986). One example would be to ask students to apply course concepts through realistic problems or scenarios (see below; see also: Suskie, 2010). For example, Crowe, Dirks, & Wenderoth (2008) developed the "Blooming Biology Tool" to help instructors align their assessments with higher order teaching activities. Nilson (2016) and Suskie (2010) note that effective MCQs must be phrased carefully to avoid accidentally steering a prepared student away from the correct response, or alternately steering an unprepared student to the correct answer. Following are suggestions for designing effective MCQs:

Strategies	Explanation	Teaching Suggestions
Make sure the statement or question is clear and concise.	Lengthy, unclear multiple choice questions can easily direct even prepared students to the incorrect answer, and produce considerable anxiety and frustration for students (Suskie, 2010).	Barbara Mills, Test Specialist with CEE suggests avoiding overlapping answers (a particular issue when numbers are the answer choices) and numerical answers that are too close (such as those distinguished only by rounding). Additionally, to avoid overlap, it's best to use mutually exclusive response options, and to include only one correct, clearly best answer.
Use consistent and clear language	Write a simple, straight-forward question or prompt. Use language consistently. Be concise. Avoid turning your content test into a test of reading comprehension. Give students the opportunity to focus on the task itself and not on puzzling through the question (Suskie, 2010).	Examples of inconsistent language usage include: alternating verb tenses within the question and or using different pronouns in the response choices from those in the question stem (see below for example).
	It's important to write items that are clearly written, precise, and accessible. To do this: use sentence structures and vocabulary that are appropriate for the audience; eliminate difficult or unclear terminology (OR undefined acronyms!); and avoid potential cultural and/or linguistic issues.	Read each stem out loud, followed by each response. *Listen* for language "bumps" and address them. Also, consider asking your TAs or another instructor to read through your questions for clarity.
Tie questions to specific learning goals for the course/unit	An intentionally aligned instructional approach "provides students opportunities to synthesize, practice, and develop increasingly complex ideas, skills, and values" (Allen, 2004, p. 40). In addition, if teaching and learning activities, including modes of instruction and assignment design, are not aligned to the goals set for students, instructors will be unable to demonstrate the excellent work in teaching that they do (Jankowski, 2017).	All assessments in a course should derive directly from the course learning outcomes. Remind students *throughout the course* to use the course outcomes as reference points for their own learning.
	Begin with the course learning outcomes to ensure that assessment activities are directly aligned to your goals for student learning. Design test questions that assess what students should be learning in your course (Handelsman, Miller, & Pfund, 2007; McKeachie & Svinicki, 2013; Suskie, 2010).	Zimmaro (2004) suggests avoiding asking about trivial information or unimportant facts, as doing so can lead to further test anxiety when students don't know what information is actually important.
Employ "stimulus-based" multiple choice questions as a way to tie the assessment to interpretive or	Nilson (2016) describes these items as a series of multiple choice questions corresponding to a realistic stimulus like a text passage, table, graph, image, equation, description of an experiment or short case example, etc. For an	When designing these types of questions, make sure to give students prior practice (in class, on homework, or on a study guide) in interpreting the types of stimuli you intend to use on the exam. But, make sure the scenarios and examples are new to students (Suskie,

applicative thinking skills.	example of this type of question, see "Example Questions" below.	2010). Additionally, the longer or more complex your stimulus is, the more questions you should include in your corresponding series.
Avoid assigning "all of the above" and/or "none of the above" options as the correct response	Most test designers (e.g., Haladyna, 2004; Nilson, 2016; Suskie, 2010) agree that when "all of the above" or "none of the above" options are the correct response, this makes it easier for students to select the correct answer without actually knowing the material. Barbara Mills also suggests avoiding choices such as "A & C" or "B & D," as these can also make it harder to distinguish between students who know the material and those who don't.	Nilson (2016) suggest using "all of the above"/"none of the above" options as distractors, as this can make a question more challenging to students, and ensure that they actually need to know the material to find the correct response. You can also use common mistakes, misconceptions, or mis-associations that students make (Suskie, 2010), or alter elements of or variables within the correct response to design distractors.
Avoid using negative phrasing, or clearly signal the negative word to students.	Using negative phrasing can confuse a student, even if they know the material, especially if they are short on time (Clegg & Cashin, 1986; Haladyna, 2004; Suskie, 2010).	For example, "Under which of the following conditions is X not true?" can be easily misread by students. A better version would be "Under which of the following conditions is X true?"
		Additionally, highlight, bold, all-cap, or underline negative words to signal to students what is being asked (e.g., Zimmaro, 2004). For example, "Which of the following countries is **not** on the UN Security Council?"

Example Questions
Examples in this section come from Nilson (2016), and this resource from the Vanderbilt University Center for Teaching.

Item responses use inconsistent language and negative phrasing

Which of the following is not true about mitochondria?
  a: They contain DNA
  b: Mitochondria made some of their own proteins
  c: They are static
  d: none of the above
 	
The negative phrasing can be easy for students to miss, especially if they are worried about time. Additionally, the phrasing for option "b" is inconsistent from the rest of the questions as it uses past tense (while the other options are phrased in the present) and it uses the word "Mitochondria" instead of the pronoun "they." A better phrasing would be:
 	
*Which of the following is **not** true about mitochondria?*
  a: They contain DNA
  b: They make some of their own proteins
  c: They are static
  d: none of the above

Example of "stimulus-based" test items
Two researchers were studying the relationship between amount of sleep each night and calories burned on an exercise bike for 42 men and women. They were interested if people who slept more had

> *more energy to use during their exercise session. They obtained a correlation of .28. With a two-tailed probability of .08, and the alpha was .10.*
>
> *What is the correct statistical null hypothesis?*
>   *a: There is no correlation between sleep and energy expended.*
>   *b: rho equals zero.**
>   *c: R equals zero.*
>   *d: rho equals r.*
>
> *What conclusions should you draw regarding the null hypothesis?*
>   *a: Reject**
>   *b: Accept*
>   *c: Cannot determine without more information. (Nilson, 2016)*
>
> For this stimulus-based question, the test designer wrote out a research scenario for a statistics class, then designed a series of questions referencing several different course concepts (i.e., the null hypothesis). (For the full set of questions related to this scenario, see Nilson, 2016)

Additional resources

- At UC Davis, instructors can contact Barbara Mills, Testing Specialist (bjmills@ucdavis.edu) in the Center for Educational Effectiveness for support in designing test questions.
- This resource was designed with the help of Kara Maloney, PhD, Assessment Lead in the Center for Educational Effectiveness (kmoloney@ucdavis.edu).

Citation

Center for Educational Effectiveness [CEE]. (2018). Writing Effective Test Questions Series. *Just-in-Time Teaching Resources.* Retrieved from https://cee.ucdavis.edu/JITT

References

Allen, M. J. (2004). Assessing academic programs in higher education. San Francisco, CA: Jossey-Bass/Anker.

Center for Research on Learning and Teaching, University of Michigan [CRLT]. (n.d.). *Framework for Designing Effective Exams*. Retrieved from http://crlt.umich.edu/olws/6/framework

Center for Research on Learning and Teaching, University of Michigan [CRLT]. (n.d.). *Writing Questions*. Retrieved from http://crlt.umich.edu/olws/6/questions

Clegg, V. L, & Cashin, W. E. (1986). Improving Multiple-Choice Tests. *IDEA Paper No. 16*. Retrieved from: www.theideacenter.org

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology. *CBE-Life Sciences Education, 7*(4), 368-381. Retrieved from http://www.lifescied.org/content/7/4/368.full

Haladyna, T. (2004). *Developing and Validating Multiple-choice Test Items.* Ebook: Ebrary, Inc.

Handelsman, J., Miller, S., & Pfund, C. (2007). *Scientific teaching.* New York, NY: Macmillan.

Jankowski, N. A. (2017). Unpacking relationships: Instruction and student outcomes. *American Council on Education*. Retrieved from http://www.acenet.edu/news-room/Documents/Unpacking-Relationships-Instruction-and-Student-Outcomes.pdf

McKeachie, W., & Svinicki, M. (2013). *McKeachie's teaching tips.* Belmont, CA: Cengage Learning.

Momsen, J. L., Long, T. M., Wyse, S. A., and Ebert-May, D. (2010) Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *Cell Biology Education, 9* (Winter), 435-440. Retrieved from http://www.lifescied.org/content/9/4/435.short

Nilson, L. B. (2016). *Teaching at its best: A research-based resource for college instructors* (4th ed.). San Francisco, CA: Jossey-Bass.

Suskie, L. (2010). *Assessing student learning: A common sense guide.* (2nd ed.). San Francisco, CA: Jossey-Bass.

Zimmaro, D. M. (2004). *Writing good multiple-choice exams.* Measurement and Evaluation Center. University of Texas at Austin. Retrieved from http://www6.cityu.edu.hk/edge/workshop/seminarseries/2010-11/Seminar03-WritingGoodMultipleChoiceExams.pdf