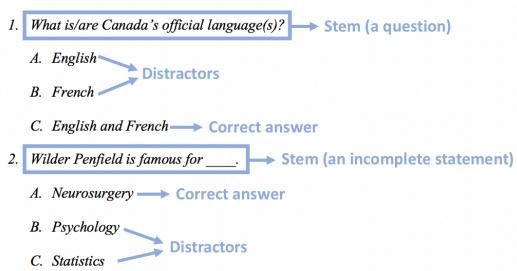# MULTIPLE CHOICE AND OBJECTIVE TESTS

A Quick Overview

What is it?

Objective tests are ones that have finite correct answers and can be easily graded without subjective judgement, even by machine.

The most common form of objective tests are multiple choice tests. Each multiple-choice test *item* presents a *stem*, which could be a question or an incomplete statement. Several possible conclusions follow—these are the choices. One is correct. The other choices are called *distractors* because they are intended to distract learners from the correct response.

Example:

1. What is/are Canada's official language(s)? → **Stem (a question)**
   A. English ⟍
   B. French ⟋ **Distractors**
   C. English and French → **Correct answer**

2. Wilder Penfield is famous for _____. → **Stem (an incomplete statement)**
   A. Neurosurgery → **Correct answer**
   B. Psychology ⟍
   C. Statistics ⟋ **Distractors**

Other types of questions used on objective tests include:

- True/false questions, which ask learners to distinguish undeniably correct and incorrect statements. True/false assessments typically involve students receiving a series of statements; students must indicate whether each is correct (true) or not (false).

- Multiple-response questions, which are similar to multiple choice questions including stems and options. However, multiple-response questions have more than one correct answer.

- Matching questions, which provide two lists of items and give guidance on how to match individual items in one column with those in the other.

Benefits

- Test all students at one time (called group administration), even if several versions of the tests are offered (Ng & Chan, 2012; Xu, Kauer, & Tupy, 2016). This makes the assessment of a large number of students an efficient process.

- Assess a larger quantity of course material than other types of tests. That's because students can complete multiple-choice and other objective questions more quickly than other types of questions (like essays). So in the time of a single test, a multiple choice test can ask more questions and, at least theoretically, cover more material (Bacon, 2003).

- Because each response has one and only one correct response, multiple choice tests can be marked efficiently and reliably. Grading equipment like Scantron can be used (Ward, 1981).

- When well-constructed, distractors can be used to identify possible causes of learners' misunderstandings, and can help with efforts to correct them (a process called *remediation*) (Butler & Roediger, 2008; Little & Bjork, 2015; Towns, 2014).

- Because the results of multiple-choice tests do not require subjective judgement, there is no rater effect and the cost of scoring is low (Haladyna, 2012).

- Multiple choice questions have uses beyond quizzes, tests, and exams. Instructors can use them to engage learners and spark classroom discussion—especially in large classes designed using active learning techniques.

Challenges

Battling common misconceptions:

- One is that multiple choice tests can only assess lower-order cognitive skills such as matching, recognizing, and simple calculations (Xu, Kauer, & Tupy, 2016). Well-constructed multiple choice tests can assess higher-order skills such as analysis, synthesis, and evaluation.

    But according to previous research, properly developed objective tests yield similar results when compared the results of other types of assessments (Bacon, 2003; Tasdemir, 2010).

- Another is that students can guess their way through the exam because one of the possible answers is the correct one (Burton, 2001; Kubinger, Holocher-Ertl, Reif, Hohensinn, & Frebort, 2010).

    But properly developed tests (a) emerge from the course objectives and (b) have well-chosen distractors (incorrect responses) that can be linked to ways of thinking about the question. As a result, the distractors look more plausible (reducing

guessing) and help detect common errors in thinking, which can later be addressed with students (Butler & Roediger, 2008; Little & Bjork, 2015; Towns, 2014).

How to use it?

Note: This section assumes that instructors will write their own test questions. In certain courses or when using certain textbooks, instructors have access to existing tests. The details provided here provide instructors with some guidance in assessing the quality of those existing tests.

1. Define the objectives of the test (Downing & Haladyna, 2006).

   Use the objectives in the syllabus (also called the course outline). Note that the action word in the objectives, such as compute, describe or define, should align with the test question.

   Example: Here are a list of objectives for a communication skills course:

   a. *Make eye contact with all group members*

   b. *Use other members' names in the conversation*

   c. *Use small talk to warm-up the conversation*

   d. *Set up the goal for the communication*

   e. *State opinions clearly*

   f. *Express disagreement or doubt without causing conflict*

   g. *Ask for repetition*

   h. *Seek for clarification or elaboration*

2. Determine the length of the test, such as one hour or two hours. This is likely to be determined by the length of the class session or the exam period specified by your institution.

3. Determine the total number of questions to ask. The rate at which students can answer questions varies widely based on the subject matter and the nature of the test material. For example, students can more easily answer questions that involve recall of facts than determining which of three choices is the correct response to an equation (which students also have to work through to choose a correct response). Table 1 provides some general guidance, though actual needs might differ from the suggestions offered in the table.

Time available	Number of simple questions	Number of long or complex questions
1 hour	50 to 60	40 to 45
90 minutes	72 to 80	65 to 70
2 hours	100 to 110	80 to 90

Table 1. General suggestions on the time needed to answer particular numbers of questions, as adapted from Ward (1981).

Example: Assume that students work on a recall question and you have estimated that they need 80 seconds to answer one question, or about .75 questions per minute. For a 1-hour (60-minute-) exam, that's 45 questions. To provide a little "breathing room" for students who might need a little extra time, you might round that down to 40 questions.

4. Set up the scoring system by choosing among these options (Lesage, Valcke, & Sabbe, 2013):

   a. Number-right scoring: Correct answers are scored with a positive value while incorrect answer and absent answers were scored with a value of zero. The test score is the sum of the scores for correct responses.

   Example: A multiple-choice test has 50 questions where each question has a value of 1. A student answered 35 questions correctly, 14 incorrectly, and did not answer 1 question. The student's score in this test should be 35.

   b. Negative marking or "Rights minus wrongs" model: Correct responses receive a positive score; students are penalized for incorrect responses. Omitted items results in no mark. The formula for calculating the penalty of an incorrect answer is:

   Penalty score per question = $\dfrac{1}{n-1}$

   Where *n* is the number of options.

   *Example: A multiple-choice test has 50 questions where each question has a value of 1. There are three options for each question: correct, incorrect, or no response. A student answered 45 correctly, 3 incorrectly, and did not respond to 2. For each wrong answer, the penalty is: 1 / (3-1) = 0.5. The student answered 3 questions incorrectly, then the penalty for the four questions should be: 3 * 0.5 = 1.5. The one omitted item results in 0. Therefore, the student's final score should be: 45 – 1.5 + 0 = 43.5.*

   One last point: this is the most frequent marking approach used.

c. Confidence-based marking (Nicol, 2007): Students should select the answer that they think is correct and rate their confidence level in their response. The follow is an example:

Confidence (Certainty degree)	High	Medium	Low	No reply
Mark if correct	3	2	1	0
Penalty if wrong	-6	2	0	0

d. Correction for guessing (Ward, 1981): Correct each student's score using the following formula:

$$S = R - \frac{W}{n - 1}$$

S is the corrected score

R is the number of questions answered correctly

W is the number of questions answered wrongly

n is the number of options

Example: A multiple-choice test has 50 questions where each question has a value of 1. There are three options for each question. A student answered 44 correctly, 6 incorrectly. The student's score for the test should be: 44 – 6 / (3-1) = 41

5. Divide the list of objectives into natural sections, which align with the sections of the course.

Example: We use the list of objectives in the previous section. The first three objectives (a, b, c) belong to "Initiate Conversations". The objectives d, e, and f constitute the section "Express opinions without causing conflict". The last two

objective g and h belong to the section "Make sure the information from the speaker is received correctly".

6. Determine the number of questions for each objective (Burton, 2006; Ward, 1981):

    a. Because students cannot always guess the correct answer, ask a minimum of two questions for each objective, more if possible.

    b. Assume that students need around 80 seconds to answer one question.

    c. Determine the total length of the test.

    d. Given the number of objectives to be assessed, the rate at which students are expected to complete the test, and the number of items to test each objective, choose a total number of questions.

    Example: You have 20 objectives for this test, and you would love to include 2 questions per objective. So in total you will have 40 questions for the whole test. Assume that students need 80 seconds to answer each question, then the total time that students need to finish answering the questions will be: 60 * 80 = 3200 seconds = 53.3 minutes. You might consider adding 15 minutes for students to go through the test again in the end. Then the total length of the test could be around 70 minutes.

7. Write the test questions.

    a. Determine the number of questions you need to write for each objective.

        i. In addition to the number of questions you intend to include on each test, multiply that number by the number of versions of the test you intend to prepare.

Example: If you intend to write three questions to test objective 5 on the test, but plan to offer three versions of the test, you will need nine questions altogether (3*3).

    ii. If you plan to prepare practice questions (always a good idea), multiply the number of questions again.

   iii. If you plan to write the exam for several years and are concerned that students might share tests after taking them, take that into account in the total number of test questions for each objective (Kun, 2015). You would need to write additional questions.

Note: This collection of test questions is called a *test pool.*

b. Write the stems (a question or an incomplete statement)) of the questions (Haladyna, 2012; Xu, Kauer, & Tupy, 2016; Ward, 1981).

    i. What to cover:

- Only address content that has been taught to the students in class.

- Avoid trivial content, over-specific and over-general content.

    ii. Make sure that the stem provides clear and necessary information about the knowledge that is being tested.

   iii. Keep the stem short (when possible 25 words maximum, fewer if possible). Avoid unrelated information in the stem to reduce the reading "load" for students.

   iv. Avoid writing stems as negative statements. These confuse students. Furthermore, some students will not notice the negative word under the examination pressure.

Example: Which of the following are not examples of good governance?

v. Avoid opinion-based stems.

Example: Who is the best singer?

vi. Write questions assessing near transfer. Near transfer refers to situations are similar to those taught in class or tested in practice questions. Most students should successfully answer these questions.

vii. Write questions assessing far transfer. Far transfer questions present students with scenarios that substantially differ from those seen in class but apply the same knowledge, skills, and abilities all the same. Far transfer is a mean of distinguishing between those students who have merely mastered the basic material and those who can apply it in novel situations.

c. Write the correct responses.

i. Only provide one correct response to each question.

ii. The correct response should be unarguably right, not subject to other factors.

Example: What is the most likely reason for students to drop courses?

A. There are too many assignments

B. Students don't like the atmosphere of the class

C. The course is not helpful for their career path

All the choices in the example above could be true, so no unequivocal correct response exists.

d. Write the distractors (DiBattista, Sinnige-Egger, & Fortuna, 2014; Little & Bjork, 2015; Rodriguez, 2005; Roediger III & Marsh, 2005; Towns, 2014; Yaman, 2011).

    i. Distractors should be unquestionably wrong answers.

    ii. Distractors should still be plausible enough so that students who don't know the correct answer might choose them. Some tips for choosing distractors:

- For conceptual ideas, choose related concepts that are similar but not correct.

  Example: The moon is a ____

  a. Star

  b. Planet

- Also for conceptual ideas, choose known misconceptions.

  Example: Issues that instructors need to consider when designing university courses include:

  a. Learning styles

  b. Objectives

- For calculations, choose ones that would result from arithmetic errors or making an error at an earlier step of the calculation.

  Example: 46 + 65 = _____.

  A. 101

  B. 111

iii. Develop as many distractors as possible but only choose two in the end. If you plan to use the same questions for several years, you could select different distractors each year to avoid cheating.

Note: Research suggests that the ideal length of multiple choice questions is 3 options: 1 correct response and 2 distractors.

e. Edit the correct response and distractors.

i. Avoid the use of "*All of the above*" because that option could encourage guessing if students determine that one or two answer are correct.

*Example you should **not** follow:*

Which of the following cities, 1) Montreal, 2) Rimouski, 3) Sherbrooke, are located in Quebec?

A. Only 1

B. Only 1 and 2

C. Only 1 and 3

D. 1, 2 and 3

ii. Avoid the use of "*None of the above*" because it is hard to tell whether the students really know the correct answer or not.

Example you should not follow: What is the standard voltage of Canada?

a. 110V

b. 120V

c. 220V

d. None of the above

iii. List options in alphabetical or numerical order.

Example you **should** *follow: Who was the first prime minister of Canada?*

a. *John Diefenbacher*

b. *John A. MacDonald*

c. *Pierre Elliott Trudeau*

Note that the prime ministers are listed in alphabetical order by last (family) name, as is the custom for organizing names in English.

Another example you **should** *follow: Which of the following is a prime number?*

a. *1*

b. *269*

c. *374*

Note that the numbers are listed in increasing numerical order.

iv. Assign the position of the correct choice randomly

Note: —if you list all options alphabetically or in numerical order, this happens naturally

v. Keep the structure and the length of options about the same (including the correct answer and distractors).

8. Prepare the test for distribution to students.

a. At the beginning of the exam, provide brief, clear instructions to students that address:

- The length of the test

- The types of questions

- How to record responses

- How to submit completed exam

- Other relevant instructions

Example:

Instructions: *This test consists of 50 multiple-choice questions. You have 2 hours to complete this test. You must use a 2B pencil to mark your chosen answers on the answer sheet. If you want to change your answer, erase the existing one thoroughly, then mark the newly chosen answer. After completing the exam, submit your answer sheet to the instructor.*

   b. When formatting the test, make sure questions and their possible responses all appear on the same page or in the same column.

   Note: If a question and responses spill onto another page or column, some students will miss them. This could lead to challenges of the exam.

9. After administering the test, but ***before*** releasing the results, review the test to make sure that the test is fair and students who keep up with their work can pass it (McCoubrie, 2004; Towns, 2014; Ward, 1981). Table 2 identifies indicators of issues and suggests the issue that should be explored.

Indicator	Possible Issue to Explore	Possible Solution(s)
• Many students did not finish the test • Nearly all of the students finished the test before the time was up	Whether sufficient time is available	• If too many students did not finish the exam, either provide more time or reduce the number of questions. • If all of the students finished early, reduce the time for the exam or increase the number of questions.
The majority of students incorrectly answer a question	Whether the question is fair	If the percentage of students correctly answering questions is low, analyze the question: • Did a large percentage of students choose one of the distractors rather than the correct response? Some confusion in the question or distractors might need to be resolved. • Is it directly linked to a course objective? If not, replace the question. • Did homework and class sessions properly prepared students to answer it? If not, fix the lesson plans.
Reliability coefficient is used to indicate the test's internal	Whether all the items are consistent in its measurement	Analyze the percentage of students answering each question correctly.

consistency is below 0.75 (.075 is acceptable for university tests)		• If too many students correctly answer too many questions, consider whether the question asks enough questions assessing far transfer (situations different from class).
Standard deviation across scores is greater or less a 10-to-12-percent range	Whether the test sufficiently differentiates among students of different ability	• If the percentage of students correctly answering questions is low, ask whether it is clear? linked to the objectives? whether the class properly prepared students to answer it? or it the test has too many far transfer questions?

Table 2. Indicators of issues with an exam and issues these indicators suggest.

REFERENCES

Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice

and short-answer questions in a marketing context. *Journal of Marketing*

Education, 25(1), 31-36.

Burton, R. F. (2001). Quantifying the effects of chance in multiple choice and true-false

tests: Question selection and guess of answers. *Assessment & Evaluation in*

Higher Education, 26(1), 41-50. doi: 10.1080/02602930020022273

Burton, R. F. (2006). Sampling knowledge and understanding: How long should a test

be? *Assessment & Evaluation in Higher Education, 31*(5), 569-582. doi:

10.1080/02602930600679589

Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and

reduces the negative effects of multiple-choice testing. *Memory & Cognition,*

36(3), 604-616.McCoubrie, P. (2004). Improving the fairness of multiple-choice

questions: a literature review. *Medical Teacher, 26*(8), 709-712. doi:

10.1080/01421590400013495

DiBattista, D., Sinnige-Egger, J., & Fortuna, G. (2014). The "noun of the above" option

in multiple-choice testing: An experimental study. *The Journal of Experimental*

Education, 82(2), 168-183.

Downing, S. M., & Haladyna, T. M. (Eds.). (2006). Handbook of test development.

Mahwah, Jew Jersey: Laurence Erlbaum Associates.

Haladyna, T. M. (2012). Developing and validating multiple-choice test items (3rd ed.).

Mahwah, NJ: Lawrence Erlbaum Associates.

Kubinger, K. D., Holocher-Ertl, S., Reif, M., Hohensinn, C. & Frebort, M. (2010). On

   minimizing guessing effects on multiple-choice items: superiority of a two

   solutions and three distractors item format to a one solution and five distractors

   item format. *International Journal of Selection and Assessment, 18*(1), 111–115.

Kun, A. I. (2015). Does the repetitive use of the same test in consecutive examination

   sessions facilitate cheating? *Assessment & Evaluation in Higher Education,*

   40(8), 1159-1181. doi: 10.1080/02602938.2014.977845

Lesage, E., Valcke, M., & Sabbe, E. (2013). Scoring methods for multiple choice

   assessment in higher education -- Is it still a matter of number right scoring or

   negative marking? *Studies in Educational Evaluation, 39*(3), 188-193. doi:

   10.1016/j.stueduc.2013.07.00

Ng, A. W. Y., & Chan, A. H. S. (2012). Evaluation of three multiple-choice assessment

   methods in a human factors engineering course. *Journal of the Chinese Institute*

   of Industrial Engineers, 29(7), 466-476. doi: 10.1080/10170669.2012.727478

Nicol, D. (2007). E-assessment by design: Using multiple-choice tests to good effect.

   Journal of Further and Higher Education, 31(1), 53-64. doi:

   10.1080/03098770601167922

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-

   analysis of 80 years of research. *Educational Measurement: Issues and Practice,*

   24(2), 3-13. doi: 10.1111/j.1745-3992.2005.00006.x

Roediger III, H. L., & Marsh, E. J. (2005). The positive and negative consequences of

   multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory,*

   and Cognition, 31(5), 1155-1159. doi: 10.1037/0278-7393.31.5.1155

Tasdemir, M. (2010). A comparison of multiple-choice tests and true-false tests used in

    evaluating student progress. *Journal of Instructional Psychology, 37*(3), 258-266.

Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple

    choice assessments. *Journal of Chemical Education, 91*, 1426-1431.

Ward, C. (1981). Preparing and using objective questions. London: Stanley Thornes.

Xu, X., Kauer, S., & Tupy, S. (2016). Multiple-choice questions: Tips for optimizing

    assessment in-seat and online. *Scholarship of Teaching and Learning in*

    Psychology, 2(2), 147-158. doi: 10.1037/stl0000062

Yaman, S. (2011). The optimal number of choices in multiple-choice tests: Some

    evidence for science and technology education. *New Educational Review, 23*(1),

    227-241.