



WWW.LEANCOMPETENCY.ORG

GUIDE TO DEVELOPING MULTIPLE CHOICE TESTS

CONTENTS

Introduction.....	3
Advantages & Disadvantages of MCTs	3
Advantages.....	3
Disadvantages	3
Development of Tests	4
Alignment to LCS Detailed Level Description.....	4
Question Design.....	5
Question Structure.....	5
Tips for constructing multiple choice questions.....	5
Types of Questions.....	6
Comparison with Examinations	8
Online Tests	8
General Guidelines When Setting Up Online Tests	8
Types of Learning Suitable for MCTs	10
Appendices	12
1. Ten Rules for Developing Tests.....	12
2. Further Information & References.....	13

INTRODUCTION

LCS accreditation requires that course participants are assessed after completing a learning programme to check that they have appropriate knowledge for a particular level. There are several assessment approaches that can be taken and multiple choice tests (MCTs) are a popular option and commonly form part of a training programme or system's assessment strategy.

This document highlights some of the advantages and disadvantages of MCTs, discusses how learning outcomes need to be aligned with level descriptors, examines question design and suggests general principles for their construction.

ADVANTAGES & DISADVANTAGES OF MCTS

ADVANTAGES

1. Versatility: they are appropriate for use in many different subject-matter areas and can be used to measure a great variety of educational objectives.
2. They are adaptable to various levels of learning outcomes, from simple recall of knowledge to more complex levels – eg ability to analyse, apply principles to new situations, comprehend concepts, interpret cause-and-effect relationships, solve problems, etc.
3. When carefully designed, they can assess higher-order thinking skills, where candidates are required to generalise, analyse and make inferences about data.
4. Multiple choice tests often require less time to administer for a given amount of material than tests requiring written responses. This results in a more comprehensive evaluation of the candidate's extent of knowledge.
5. Efficiency can be created by using online testing software.

DISADVANTAGES

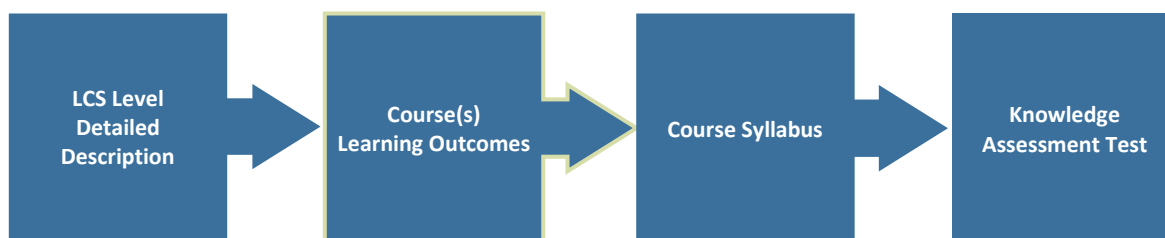
1. It is argued by some that limited types of knowledge that can be assessed by MCTs and that they are best used for testing well-defined or lower-order skills, whereas problem-solving and higher order reasoning skills are better assessed through short-answer and essay type tests.
2. Where candidates have some knowledge of a question, they receive no credit for knowing that information if they select the wrong answer.
3. A candidate who is not capable of answering a question can simply select a random answer and still have a chance of receiving a mark for it (“multiple guess”).
4. Poorly phrased questions lead to ambiguously and may cause confusion.

DEVELOPMENT OF TESTS

ALIGNMENT TO LCS DETAILED LEVEL DESCRIPTION

Since one of the main the purposes of an assessment is to ensure that learning outcomes have been achieved, there needs to be a clear linkage between the test and the detailed description of a particular LCS level. The test is effectively checking that an individual possesses the right LCS level knowledge.

Therefore, there should be a clear connection between the LCS level detailed description, course learning outcomes, course syllabus and the test.



A useful check to ensure that there is alignment is to produce an **assessment - learning outcomes alignment matrix**:

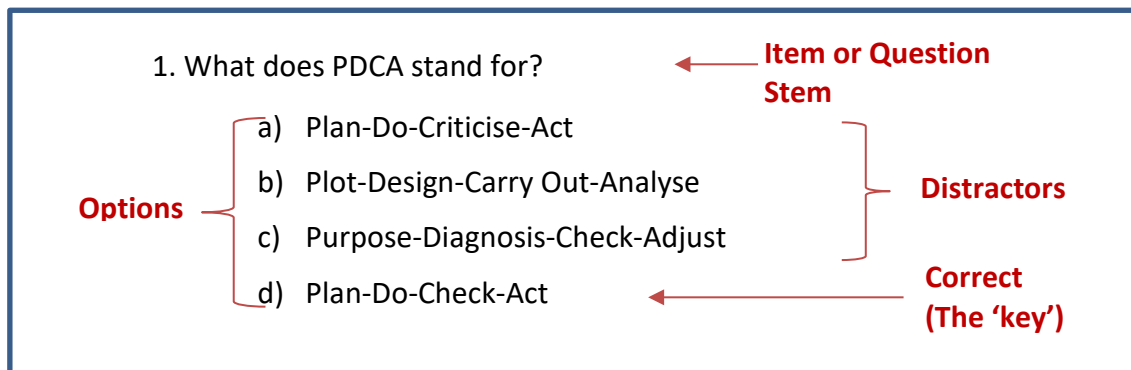
Assessment (eg test questions, essay, interview etc)	Learning Outcomes (listed below)						
	1	2	3	4	5	6	etc
List Womack & Jones five lean principles	X F		X P				
etc							
Learning Outcomes 1. Know the Five Lean Principles 2. etc							

Test questions, for example, are listed in the first column and a tick placed in the column of the particular learning outcome to which the question relates. The table can be tailored to provide different information – for example, it can be shown whether the question fully (F) or partially (P) addresses a particular learning outcome.

QUESTION DESIGN

QUESTION STRUCTURE

Multiple choice questions (or 'items') are made up of problem (referred to as the 'stem')



and a list of suggested solutions or options. The stem could be either a question or an incomplete statement and the list of options contains one correct or best alternative answer, plus several incorrect or inferior options (referred to as 'distractors').

Good multiple-choice questions are generally more difficult and time-consuming to write than other types of tests and devising plausible distractors requires some skill, though this may be developed through practice and experience. It is also important to word the stem so that there is a single correct answer.

Stems can be statements or questions, or be in a negative voice ("which of the following is not one of the 7 wastes?") or a conditional voice ("if you wanted to use a technique that speeded up your equipment changeover, you would select:").

TIPS FOR CONSTRUCTING MULTIPLE CHOICE QUESTIONS

Writing stems

- Present a single, definite statement to be completed or answered by one of the several given choices.
- Avoid unnecessary and irrelevant material.
- Use clear, straightforward language in the stem of the item. Questions that are constructed using complex wording may become a test of reading comprehension rather than an assessment of whether the candidate knows the subject matter.
- Use negatives sparingly. If negatives must be used, capitalize, underscore, embolden or highlight.
- Put as much of the question in the stem as possible, rather than duplicating material in each of the options.

Writing distractors

- For single response MCQs, ensure that there is only one correct response. Use only plausible and attractive alternatives as distractors.

- Avoid giving clues to the correct answer.
- Avoid the choices “All of the above” and “None of the above”. If you do include them, make sure that they appear as correct answers some of the time.
- Distracters based on common student errors or misconceptions are very effective.
- Correct statements that do not answer the question are often strong distracters.
- Avoid using ALWAYS and NEVER in the stem
- Do not create distracters that are so close to the correct answer that they may confuse students who really know the answer to the question.
- Provide a sufficient number of distracters. Four is probably the optimum number, though three or five are acceptable.

TYPES OF QUESTIONS

The table below provides examples of different types of questions that can be used in a test.

Type of question	Notes, answer choices	Example
Calculated Numeric Response	This question resembles a fill-in-the-blank question. The user enters a number to complete a statement. The correct answer can be a specific number or within a range of numbers. The answer must be numeric, not alphanumeric.	Q: 1+2=? A: 3
Either/Or questions	Agree/disagree, yes/no, right/wrong, true/false	The first chief engineer of Toyota was Taiichi Ohno. (answer: yes or true or right)
Essay	Student writes in a response to a question. This has to be marked/graded manually.	Describe the main benefits of applying lean in your company (500 words)
File response	Student attaches a file as the answer (eg word document, excel spreadsheet, photo etc). This has to be marked/graded manually.	Completed Kobayashi 20 Keys chart
Fill in multiple blanks	Add question text that includes several variables in square brackets [x] corresponding to locations in the text where the fill in the blank fields should appear. Variables must be unique and cannot be reused	The original company name or Toyota was [toyoda] and it made [looms] before it made cars
Fill in the blank	Between 4 and 20 correct answers can be given	Name one of the seven wastes
Hot spot	Users indicate the answer by marking a specific point on an image (that we upload). A range of pixel coordinates is used to define the correct answer. Hot Spot refers to the area of an image that, when selected, yields a correct answer.	Picture of a factory floor with some obvious quality or 5S problems that the user has to click on
Jumbled sentence	Create a question text that includes variables in square brackets [x] corresponding to locations in	The best way to remove [waste] is to first [map] a value [stream]

Type of question	Notes, answer choices	Example
	<p>the text where the drop-down list of answers should appear.</p> <p>In the example on the right, for the student, all 5 words appear in a drop down menu at each set of brackets and the student has to select one.</p>	and [brainstorm] improvements for the [future] state.
Matching	Up to 20 matches allowed	<p>Match the car company with the word.</p> <p>Car comp: Toyota, Ford, over, Renault.</p> <p>Word: Lean, Mass, Bust, French</p>
Multiple answer	Between 4 and 20 possible answers can be given, any number of which can be correct	<p>Which of these is part of 5S workplace organisation?</p> <p>Answer: ✓ Sort, Simplify, ✓ Standardise, Swear, ✓ Shine</p>
Multiple choice	Between 4 and 20 possible answers can be given, only 1 of which can be correct	<p>Who won the FA cup in 1927?</p> <p>Choices: Cardiff, Swansea, Arsenal, Fulham</p>
Opinion scale/Likert	<p>Question type based on a rating scale designed to measure attitudes or reactions.</p> <p>The answers of these types of questions have a 5 or 6 point scale; there is usually a 'don't know' or 'not applicable' option.</p> <p>Eg strongly agree to strongly disagree</p> <p>Very interested to Not very interested</p>	<p>Toyota is the leanest car company in the world</p> <p>Strongly agree</p> <p>Agree</p> <p>Neither agree nor disagree</p> <p>Disagree</p> <p>Strongly disagree</p>
Ordering	Between 4 and 20 possible items can be presented	<p>Q: What is the correct order of the 5 principles?</p> <p>A:</p> <p>5 Perfection</p> <p>1 Value</p> <p>2 Value stream</p> <p>3 Pull</p> <p>4 flow</p>
Quiz bowl	<p>Quiz Bowl questions are a way to add fun and creativity to self-assessments. The user is shown the answer and responds by entering the correct question into a text box. An answer must include a phrase and a question word, such as who, what, or where, to be marked as correct. For example, the question may be "The person who invented the cotton gin", with the answer being "Who is Eli Whitney?" Partial credit may be given if the question word is not included in the answer.</p>	<p>Q: The first chief engineer of Toyota and father of TPS</p> <p>A: Who is Taiichi Ohno?</p>
Short answer	Similar to essay questions though the length of the answer can be controlled. This has to be manually graded.	Q: Briefly describe the seven wastes

Type of question	Notes, answer choices	Example
True/false	User has to decide whether a statement is true or false	Q: Deming was big in quality. A: True

COMPARISON WITH EXAMINATIONS

An alternative to MCTs is the **descriptive examination** and the relative advantages of each are shown below:

Task	Descriptive Questions	Multiple Choice
Setting the exam	Easy	Difficult
Marking Task	Difficult	Easy
Marking Time	Long	Short
Grade Consistency	Varies	High

ONLINE TESTS

MCT's can be undertaken physically or online. In the former, the test is supervised by an invigilator in a classroom based situation, where participants generally 'tick boxes' on a test sheet, which is subsequently marked. Online tests use assessment software (or a part of a learning management system), require a PC and internet access and can be taken in a classroom situation or 'off site' – for example, in the work place or at home.

Online testing offers advantages in terms of speed, flexibility, control and cost effectiveness.

To gain the maximum benefits of online MCTs, participants should be able to undertake them in their own time, though with some constraints. They should therefore be positioned as 'open book' – ie candidates can reference course and other material during the test.

Appendix 1 contains one expert's '10 rules for developing tests'.

GENERAL GUIDELINES WHEN SETTING UP ONLINE TESTS

- **Open or closed book:** tests can be positioned as either, though question complexity and the pass mark may need to be adjusted accordingly. If tests are taken in a test window or in a non-invigilated setting, then they automatically become open book.
- **Number of questions:** the number of questions is related to the complexity of the questions, though for LCS level 1 the expectation is for reasonably straightforward questions – eg multiple choice, select an answer from four options. Note that a typical 1b/c test would have multiple choice 40 - 50 questions. There is a range of different types of questions available.

- **Question bank:** the test questions can be selected from a question bank, which should contain at least double the amount of questions used in the test. This should be refreshed periodically, with new questions added reflecting course topic changes. Online testing systems allow questions to be selected at random from the question bank for each candidate and there is often a useful facility for a test to have fixed and random questions - for example, 25 questions that all candidates get and 25 given at random from the question bank.
- **Question format.** Use a range of types of questions (as outlined above).
- **Question marking;** the questions in a test can:
 - All have the same mark (this is recommended)
 - Have different marks, depending on question complexity
 - Have partial marking – that is, the candidate receives some marks if he/she get the question partially correct.
 - Have negative marking – that is, the candidate has marks taken off, if the answer is incorrect.
- **Piloting:** this is recommended as a general sense check and to ensure it is technically sound.
- **Test window length** (if the test is separated from the course): generally, two weeks. The time between the end of the training and opening of the window needs to be a balance of keeping the material fresh in learners' minds and giving them time to prepare.
- **Time allowance:** An allowance of one minute per question is recommended for a typical four option multiple choice question. It could be longer for candidates whose native language differs from the test questions language, or for special needs, such as dyslexia. Once the time limit is reached, the test should automatically end and the candidate's mark submitted.
- **One sitting, or pause and resume;** standard practice is to complete the test in one sitting.
- **Question ordering;** standard practice is for questions to appear in random order, so candidates do not receive questions in the same order, or do not receive the exact set of questions.
- **Answer feedback:** After submitting an answer to a question, the candidate should not be informed whether he/she got the question correct before the next question appears. Some systems allow candidates to review all their responses before they finally submit, which is reasonable (as long as it is within the time limit)
- **Pass mark;** 66% is the standard minimum for MCTs, (where there is no negative marking), though some LCS organisations pitch their pass marks as high as 80% to drive a high standard. The pass mark is a function of several variables, including marking policy, question difficulty and time allowance.

- **Number of attempts;** the standard practice is two attempts. Failing twice should require additional training or support based on an analysis of where he/she went wrong.
- **Review capability and back-tracking;** standard practice is to permit learners to review the test before finally submitting
- **Communication of test results;** learners can be informed directly of their result by the system or these can go to the test manager for distribution at a time of their choosing. Test systems allow for a range of feedback options and it is acceptable for candidates to be given a list of the answers that were incorrect, though they should not be informed on the correct answer. Tailored messages can often be given.
- **Analysis:** Most test systems allow the test manager to analyse the results of tests – for example, the average mark per question, average mark per test and reports on candidate’s performance can be generated.
- Candidates can be asked to ‘sign’ (or tick box) a ‘declaration form’ at the start of the test, confirming that they are taking the test on their own, it is their own work, without any assistance from a third party etc. This makes any transgression a clear breach.

TYPES OF LEARNING SUITABLE FOR MCTS

Bloom taxonomy¹ of educational objectives is a useful framework for categorising different types of questions. He listed six ways students can demonstrate their learning, through: Knowledge

1. Comprehension
2. Application
3. Analysis
4. Synthesis
5. Evaluation

Bloom’s Taxonomy is a simply is hierarchical system of categorising and organising understanding: each level is subsumed by the higher levels. In other words, a candidate functioning at the "application" level has also mastered the material at the "knowledge" and "comprehension" levels.

Some educationalists believe that MCTs are only useful for examining the first three or four levels of learning. However, others suggest that all six levels can be tested using MCTs.

Competence	Skills demonstrated
Knowledge	<ul style="list-style-type: none"> • Recall of information

¹ Bloom’s Taxonomy is a classification of learning objectives within education proposed in 1956, in *Taxonomy of educational objectives: the classification of educational goals*

Competence	Skills demonstrated
Comprehension	<ul style="list-style-type: none"> • Knowledge of facts, dates, events, places Question words: list, define, label, describe, name 6. Interpretation of information in one's own words 7. Grasping meaning Question words: interpret, discuss, predict, summarize, classify
Application	<ul style="list-style-type: none"> • Application of methods, theories, concepts to new situations Question words: apply, demonstrate, show, relate
Analysis	<ul style="list-style-type: none"> • Identification of patterns • Recognition of components and their relationships Question words: analyse, arrange, order, explain, connect, infer, compare, categorize
Synthesis	<ul style="list-style-type: none"> • Generalize from given knowledge • Use old ideas to create new ones • Organise and relate knowledge from several areas • Draw conclusions, predict Question words: integrate, modify, invent, design, compose, plan, formulate, arrange
Evaluation	<ul style="list-style-type: none"> • Make judgements • Assess value of ideas, theories • Compare and discriminate between ideas • Evaluate data Question words: appraise, judge, evaluate, defend

Lower level tests would usually contain a high proportion of knowledge and comprehension questions, though should also include some higher level questions. The more advanced the test, then more questions would be expected to be centred on analysis, synthesis and evaluation.

APPENDICES

1. TEN RULES FOR DEVELOPING TESTS

By Connie Malamed

http://thelearningcoach.com/elearning_design/rules-for-multiple-choice-questions/

Rule #1: Test comprehension and critical thinking, not just recall

Multiple choice questions are criticized for testing the superficial recall of knowledge. You can go beyond this by asking learners to interpret facts, evaluate situations, explain cause and effect, make inferences, and predict results.

Rule #2: Use simple sentence structure and precise wording

Write test questions in a simple structure that is easy to understand. And try to be as accurate as possible in your word choices. Words can have many meanings depending on colloquial usage and context.

Rule #3: Place most of the words in the question stem

If you're using a question stem, rather than an entire question, ensure that most of the words are in the stem. This way, the answer options can be short, making them less confusing and more legible.

Rule #4: Make all distractors plausible

All of the wrong answer choices should be completely reasonable. This can be very hard to accomplish, but avoid throwing in those give-away distractors as it detracts from the test's validity. If you're really stuck, get help from your friendly SME. (BTW, this word can also be spelled as "distracter.")

Rule #5: Keep all answer choices the same length

This can be difficult to achieve, but expert test-takers can use answer length as a hint to the correct answer. Often the longest answer is the correct one. When I can't get all four answers to the same length, I use two short and two long.

Rule #6: Avoid double negatives

No big news here, right? Don't use combinations of these words in the same question: not, no, nor, the -un prefix, etc. For example, this type of question could confuse test-takers: 'Which of the following comments would NOT be unwelcome in a work situation?' Flip it around and write it in the positive form: 'Which of the following comments are acceptable in a work situation?'

Rule #7: Mix up the order of the correct answers

Make sure that most of your correct answers aren't in the "b" and "c" positions, which can often happen. Keep correct answers in random positions and don't let them fall into a pattern that can be detected. When your test is written, go through and reorder where the correct answers are placed, if necessary.

Rule #8: Keep the number of options consistent

Did you ever have to convince a SME that he or she can't have answer choices that go to 'h' in one question and 'c' in the next? It's something of a user interface issue. Making the number of options consistent from question to question helps learners know what to expect. Research doesn't seem to agree on whether 3 or 4 or 5 options is best. Personally, I like to use 4 options. It feels fair.

Rule #9: Avoid tricking test-takers

As faulty as they are, tests exist to measure knowledge. Never use questions or answer options that could trick a learner. If a question or its options can be interpreted in two ways or if the difference between options is too subtle, then find a way to rewrite it.

Rule #10: Use 'All of the Above' and 'None of the Above' with caution

I hate this rule because when you run out of distractors, All of the Above and None of the Above can come in handy. But they may not promote good instruction. Here's why. All of the Above can be an obvious give-away answer when it's not used consistently. Also, the All of the Above option can encourage guessing if the learner thinks one or two answers are correct. In addition, the downside to None of the Above is that you can't tell if the learner really knew the correct answer.

2. FURTHER INFORMATION & REFERENCES

SlideShare presentation: http://www.slideshare.net/jhunarar/multiple-choice-tests?from_search=3

Google "developing multiple choice questions" will reveal many sites dedicated to the topic, eg <http://thelearningcoach.com/>