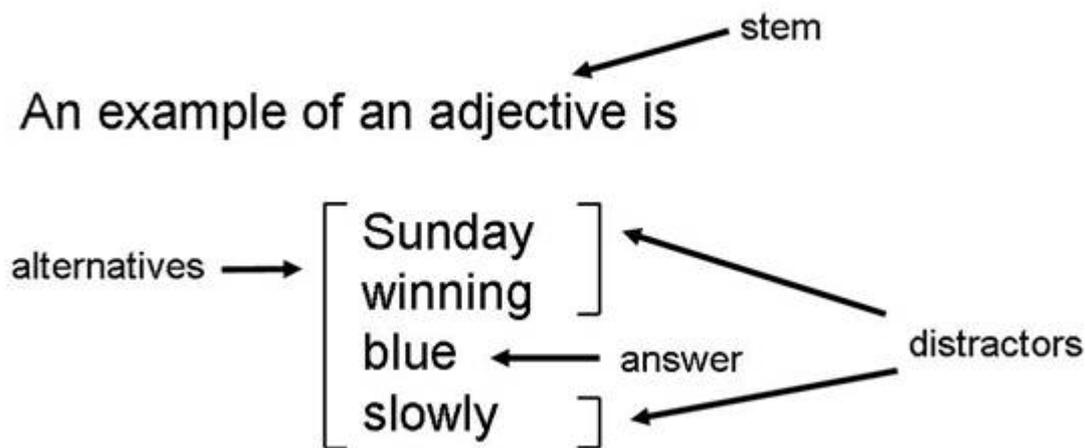


Designing Multiple Choice Questions for assessing understanding

Since a whole class needs to respond and be scored, multiple choice questions (MCQs) make for a particularly convenient method of assessment. MCQs already have some element of chance built in, since a student can select the correct answer because they know it or, occasionally, simply by guessing. This is sometimes seen as a drawback of multiple choice questions, but when teaching with games it can be positively helpful.

Every multiple choice question has a standard type of structure



MCQs have sometimes had bad press. Some might say they are limited in what they can assess, but really this is more of a reflection of how many poorly designed MCQs can be found. In reality, MCQ's greatest limitation is the imagination of whoever is designing them, since they can assess almost all levels of educational achievement. In fact, Bloom identified six levels of intellectual understanding (Anderson et al., 2001; Bloom, 1956), and all of these can be assessed using MCQ's except creating (which we'll deal with later):

- Analysing
- Evaluating
- Creating
- Applying
- Understanding
- Remembering

Each of Bloom's levels can be related to a set of verbs. Usually, the verb in the question can tell you its level – but it's not foolproof – and you can often find the same word assessing different levels. But this verb table can be really handy – although more for prompting a good range of questions, rather than for agonising over which level is being assessed!

<u>Remembering</u>	define, describe, locate, identify, know, label, list, match, name, outline, recall, recognize, reproduce, select, state
<u>Understanding</u>	comprehend, restate, convert, defend, distinguish, estimates, explain, extend, give example, interpret, paraphrase, rewrite, summarize, translate, interrelate, interpret
<u>Applying</u>	apply, change, compute, construct, generalize, dramatize, demonstrate, discover, manipulate, modify, operate, predict, prepare, produce, choose, relate, show, solve, use.
<u>Analysing</u>	analyze, break down, compare, classify, point out, categorize, contrast, diagram, differentiate, subdivide, discriminate, distinguish, identify, illustrate, infer, relate, select, separate, prioritize.
Creating	categorize, combine, compile, compose, create, devise, design, explain, generate, modify, organize, plan, rearrange, construct, produce, plan, reconstruct, relate, reorganize, revise, rewrite, summarize, tell, write, hypothesise, originate, develop, invent
<u>Evaluating</u>	appraise, compare, conclude, contrast, criticize, critique, defend, describe, discriminate, evaluate, explain, interpret, justify, relate, weight, summarize, support, consider, recommend.

To see these verbs in action, here are some examples of MCQ's for each level.

Remembering

This is the lowest type of intellectual understanding but an important one that can support all the others. This includes the recall of facts such as events, persons, terms, definitions, dates, principles and theories. For example:

Which king signed the Magna Carta in 1215?

- Richard II
- Henry I
- John
- Henry IV

What is the scientific name for the wind pipe?

- Air sac
- Trachea
- Bronchi
- Alveoli

Understanding

This includes understanding the meaning of information, demonstrated by restating it in other ways, or by interpreting, explaining or summarising the information. An MCQ that would assess this type of learning might be:

Alliteration means

- The repetition of similar vowel sounds
- The repetition of the same consonant
- The repetition of the same rhythm
- The repetition of the same metaphor

Melting point is the temperature at which

- A solid becomes a gas
- A liquid becomes a gas
- Solid becomes a liquid
- Gas becomes a solid

Applying

This involves applying rules, methods and principles to new problems or situations. This can include classifying something as a specific example.

When petroleum burns, this type of reaction can be called:

- Exothermic
- Metathermic
- Endothermic
- Electrolytic

Which of the following is correct:

- It's on the dogs' leg
- Its on the dogs' leg
- It's on the dog's leg
- Its on the dog's leg

And this type of question is very common in mathematics:

A circle of radius 2 m has a perimeter of:

- 3.14 m
- 4.00 m
- 6.28 m
- 12.56 m

Analysing

'Analysing' means identifying the organization and patterns within a system, examining and breaking up information, sometime to identify causes or trends.

For example, look at the following table:

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Week 1	2	2	5	4	6	4	6
Week 2	6	7	8	7	6	7	5
Week 3	2	4	6	4	5	4	6

Which is true?

- More rain fell on Thursdays than Saturdays
- More rain fell in Week 1 than Week 3
- More rain fell on Sunday than Thursday
- More rain fell in Week 2 than fell in Week 1 and 3 together

Or:

In "Romeo and Juliet", how might the relationship between the Friar and Romeo be characterised?

- A close friend to Romeo, but unreliable.
- A restraining influence on Romeo, but accidentally contributes to his death
- A bad influence who does not care about Romeo and contributes to his death
- An untrustworthy person who Romeo mistakes as a friend.

A scientific example of analysing might be:

A new organism is found that has a 3 part-body, 3 pairs of legs, and wings. Should it be classified as:

- An arachnid
- An annelid
- A crustacean
- An insect

Evaluating

'Evaluating' means using evidence and reasoned argument to make judgements about information, often against a set of criteria.

Stackable chairs are best made out of:

- Acrylic
- HDPE
- Expanded polystyrene
- Polystyrene

In this case, the students can apply what they know about the properties of plastics (or use information provided by the teacher) to dismiss acrylic and polystyrene as too brittle, expanded polystyrene as too weak, and evaluate high-density polythene (HDPE) as most suitable because of its hardness and toughness. Evaluation is similar to analysis, but often focuses on what's best rather than what is correct. In the next scientific example, the learner is provided with information as in the rain example above but, while all alternatives are possible, only one should be judged as best:

A scientist received the following temperature measurements from a space probe just as it passed through a comet's tail: 92, 85, 90, 9, 91, 88, 90, 89, 91, 90.

What do you think should be done with data to calculate the average?

- Ignore it and take the average of the other measurements, but include it when estimating confidence in the result.
- Ignore it and take the average of the other measurements, report it occurred but do not include it when estimating confidence in the result. Include it – and take the average of all the measurements
- Throw away the data as unreliable – no average can be safely determined

Some tips on designing MCQs

Generally:

- make questions and all the alternatives clear, concise, credible and comparable in length.
- make sure the answer appears in each position roughly the same number of times (but not exactly, because that can help give the answer away too).

That might seem simple but there's plenty of room for mistakes, as illustrated in the following semi-humorous examples (correct answers marked with *):

If the stem only makes grammatical sense with one alternative, the question is assessing an

- Memory ability
- Ability to apply basic grammar*
- Students' recall ability
- Students' understanding

In a multiple choice test, when are negative questions not confusing?

- Sometimes

- Never*
- Always
- Not never

How can you work out the answer by not even reading the alternatives?

- Pick the first one and/or most complicated, because the answer is foremost in the teacher's mind and appears most carefully thought out*
- Pick the shortest
- Pick the third
- Pick a simple answer

How do ridiculous alternatives reduce the power of an MCQ to assess learning?

- By capturing brain waves
- By attracting power-sucking aliens
- By giving power to the state
- By increasing the likelihood of a correct guess*

(Alternatives can be nonsense ones, but they have to look credible)

In what ways can more than one possible correct answer cause difficulty, if the learners can only signal one response?

- By confusing the learner*
- By making it difficult to score*
- By causing disagreement about which is more correct than the other*
- All of the above*