

# MCQs

[Pre-reading for the MCQ Workshop, University of Southampton, 23 September 2019.  
Edited extracts from Gilbert & Gale (2008), and appendices from the EASiHE (2010) and  
REAQ (2009) projects.]

During the design of a module or course, attention is given to constructing the assessment(s) which, when successfully answered by the student, demonstrates mastery of the subject matter and achievement of all of the intended student competences (also known as intended learning outcomes, competences) of the module. In most contexts, the student test is a natural part of the training or educational experience, particularly where some form of certification is required.

Ideally, every competence (top-level and enabling) will have or can be matched with a test item or assessment question which, when passed by the student, indicates successful achievement of that competence. Table 1 provides a list of the various types of test item, and identifies those that can be easily administered and scored by computer.

**Table 1. Types of test item**

<b>Computer-scorable</b>	Yes/no
	Multiple choice
	Arrange, order, or rank items in a list
	Match items in one list with those in another
	Fill-in-the-blank or provide one word or number
<b>Instructor-scorable</b>	Free response (short answer)
	Free response (essay)
	Observation of performance

## **Assessment**

### ***Formative and summative assessment***

Test items may be used in one of two distinctive ways. On the one hand, test items could be used to support the student during learning, and in this case the test items would be assembled into a formative assessment. On the other hand, the test items could be used to check that the student has indeed achieved the competences prior to being credited with the lesson, unit, or programme concerned. In this case, test items would be assembled into a summative assessment.

Another way of looking at the distinction between formative and summative assessment is to ask whether it is intended that feedback to the student on their performance in the assessment supports their learning of the topics in question. An assessment may be deemed summative if feedback from that assessment is either never returned to the student, is delayed for such a period of time as to be of little use when it is eventually returned, or is returned in such a reduced form that it cannot in fact directly assist in the learning and teaching process. An assessment may be considered formative if feedback is given to the student on their performance and is such as to materially assist their studies – that is, feedback is prompt, contingent, and specific.

The learning designer needs to take particular care to properly construct the feedback required for test items which are intended to be formative. Conversely, summative test items may involve little or no feedback.

### ***Diagnostic or routing assessment***

There is a third way in which test items might be developed and used in a module, and that is for the diagnosis of achievement of prerequisite competences and the routing of the student to appropriate lessons or units of study. Such test items would have more of the character of summative rather than formative assessment.

## **Constructing an MCQ assessment**

### ***Domain sampling***

Each question item is modest in itself. Many items will be needed, so the domain of the assessment must be clearly identified and then systematically sampled. The domain to be sampled might at first be thought to be subject matter content, but in fact it is the top level and enabling competences which are sampled. Of course, for a well-designed module, every competence has corresponding content, and every item of content is in service of a corresponding competence, so the distinction is not often important in domain sampling.

### ***Item structure***

An MCQ test item consists of a stem, a correct answer, and a number of distractors. Table 2 provides some guidelines on MCQ item construction.

**Table 2. Guidelines on basic MCQ item construction**

<b>Stem</b>	Text, diagrams, etc, specify some facts, a theory, a scenario, possible events or outcomes, etc. The stem may be quite lengthy, so that a number of questions can be asked about it.
<b>Correct answer</b>	Free of extraneous clues.
<b>Distractors</b>	<p>Free of extraneous clues.</p> <p>The incorrect options should be phrased so that they are as attractive as the right answer to someone without the required understanding, but are not so esoteric as to confuse someone who does have the required understanding.</p> <p>As many as possible. Three at the least, four preferable.</p>

## ***Evaluating item and test quality***

One of the strengths of MCQ tests is their compatibility with the use of statistics to examine and tune their performance and effectiveness.

### ***Difficulty level***

The difficulty level of an item is the proportion of students taking the test who answered the item correctly, reported as a value, “p”, between 0.0 (very difficult, no one answered correctly) and 1.0 (very easy, everyone answered correctly). Items with a range of difficulty levels are needed in a MCQ test.

For a summative MCQ, most items (perhaps 70%) should show difficulties between 0.4 and 0.6, and the rest could show difficulties as low as 0.85, and as high as 0.15. The reason for this is to enhance discrimination between students in the upper and lower quartiles. A “good” item is not judged on its difficulty alone.

For a formative MCQ, student motivation may be an issue. We use the phrase “reluctant learner” here, and suggest that such learners may benefit more from MCQ items with low difficulty, and necessarily more of them, perhaps between 0.85 and 0.95. Attention here focuses on the feedback given for a correct answer and how that supports further learning. In earlier times, such items were the core of “programmed learning” on a somewhat linear path through a curriculum. On the other hand, highly engaged learners may benefit from items with high difficulty, perhaps between 0.1 and 0.2. The focus here would be on the feedback for incorrect answers and corresponding support for more diverse paths through a curriculum.

### ***Distractor attractiveness***

Each of the distractors should be roughly equally attractive, and a histogram or count of the number of students choosing each distractor should show a uniform distribution. For example, it would be a poor MCQ item if it had a distractor which failed to be the choice of any student.

### ***Item – test correlation***

The most important measure of a "good" item is its correlation with the student's overall test mark. That is, a student who did well on the assessment as a whole should tend to get the item right, and a student who did poorly overall should tend to get the item wrong. The correlation is usually measured using the point-biserial coefficient (identical to the Pearson product-moment  $r$  when one of the variables is dichotomous), and is reported as a value, " $r$ ", ranging from 1.0 (perfect positive relationship, every student who answered correctly had the highest total test marks) through 0.0 (no relationship, the total test marks of students who answered correctly were no different from the marks of students who answered incorrectly) to -1.0 (perfect negative relationship, students who answered correctly had the lowest total test marks).

Another approach is to calculate the Student's  $t$  between the mean mark of those who passed and those who failed the item, but this is not recommended because of the inability to generalise to the case when a test item can yield a mark and not just a "0" or "1".

### ***Test psychometrics***

Psychometrics is the science of measuring cognitive and psychomotor capability and skill. There are established criteria for evaluating the quality of a measurement or test of knowledge, skills, or competencies (as well as personality traits), the most important being test reliability and test validity. These are summarised in Table 3.

**Table 3. Psychometric evaluation of tests**

<b>Psychometric consideration</b>	<b>Characteristic</b>	<b>Measure</b>	<b>Ideal value</b>
<b>Reliability</b>	Internal consistency	Correlation between items	0.8
	Alternative forms	Correlation between alternates	0.75
	Test-retest	Correlation between instances	0.7
	Inter-rater	Correlation between scorers	1.0
<b>Validity</b>	Content	Inspection of declared aims & actual content	n/a
	Predictive	Correlation with outcome	0.7
	Construct	Correlation with other measures of the same skill or competence	0.6

Test reliability takes a number of forms. One of the most useful, and one of the easiest to measure, is the internal consistency of the test items. A popular formula for calculating internal consistency is known as Kuder-Richardson formula 20, Table 4.

Table 4. KR-20 formula for internal consistency

$$KR-20 = \frac{k}{k-1} \left( 1 - \frac{\sum pq}{s^2} \right)$$

k = number of items making up the assessment  
 s = standard deviation of overall assessment raw scores  
 p = proportion of students answering an item correctly  
 q = 1 - p

Another of the strengths of computer-scorable tests is that their inter-rater reliability measure is by definition 1.0, ie a perfect correlation between the marks returned by one rater or scorer (a computer) and the marks returned for the same set of answers by another rater or scorer (a different computer).

Test validity takes three forms, known as content or face validity, predictive validity, and construct validity. Establishing the specific validity of a student test is rarely undertaken, perhaps because the test is not usually stable over time, and the cost of demonstrating or achieving adequate validity could well be equal to the cost of the entire learning unit or programme of which it is a part. Nevertheless, where the indicators for technology-based learning and teaching are positive (such as large student numbers, statutory licensing is involved, an assured quality or consistency of achievement is required, the subject matter content or procedures involve hazard, etc) some effort at establishing test validity is indicated as well.

### ***Factors affecting test psychometric statistics***

In general, the more items in a test, the higher the resulting test reliability. Almost any test can be improved by increasing its length (ie the number of questions).

In general, a test of capability or knowledge may proceed either by giving students relatively ample time to answer a relatively limited number of relatively difficult items, called a “power” test, or by giving relatively limited time to answer a relatively large number of relatively easy items, called a “speed” test. Psychometric interpretations of the usual statistical quality indicators are valid for power tests, but not for speed tests. Speed tests may be appropriate where the module competences emphasise the ability of students to quickly recall large numbers of facts or to quickly carry out a specified procedure.

### ***Test re-marking***

Another strength of computer-scorable tests is the facility to remove poor items from the test and re-mark it. The psychometric indicators mentioned earlier allow poor items to be identified, and the overall test internal consistency can be re-computed following their removal. The process is to remove the item with the lowest point-biserial, and re-compute the test total marks, observing the increase in value of the KR-20 statistic. The process can be iterated prudently.

## Guessing

Some developers and lecturers (but not students so much) are wary of computer-scorable assessments because of the perceived ease of guessing the right answers. The argument is that completely ignorant students can score 25% (4 distractors) or 20% (5 distractors) by chance.

A correction formula can be applied to allow for guessing or random responses. The formula is to subtract  $W / (Opt - 1)$  from the number of correct answers obtained by the student, where  $W$  = number of wrong answers given by the student, and  $Opt$  = number of options or distractors per question.

If a "correction formula" is applied, ie students are penalised for incorrect answers, undesirable psychological factors are introduced into the assessment, since some students are more willing to take risks than others. "Shyer" or more risk-averse students who may well know the answer, or may have a good chance of selecting the right answer, may be put off.

There is in fact no need to penalise students for guessing, and no need to apply a correction, if students are simply instructed to attempt every question.

The argument that monkeys, ie ignorant students, get 25% on a MCQ test is a simplification which does not stand up to close scrutiny. Students can get 25% in an essay without understanding very much and dressing up what they do understand in attractive prose. We may imagine that students bring some desirable intelligence and knowledge to their answers on a MCQ test, even when unsure of the correct answer. The marks distribution will probably be adjusted later anyway using a statistical technique such as scaling.

## Scaling

There is no such thing as a student's "real" mark, since an MCQ assessment could have predominantly easy items or predominantly difficult items. A student's mark is relative to both (a) other students, in this class or in previous classes (norm-referenced), and (b) some standard of achievement or performance specified by the course competences (criterion-referenced). Lecturers should have no concerns about scaling a set of marks. Their only concern should be identifying an appropriate class average and an appropriate spread of marks – that is, deciding how many students fail, and how many get the top classification, in light of their experience and understanding of the particular cohort under consideration, previous cohorts, and institutional standards for competence achievement.

Scaling, that is, adjusting the marks distribution, is simply done by applying a formula to every score, such as shown in Table 5. Scaling a set of marks transforms the test results to have a given mean and a given standard deviation. The choice of the parameters "x" and "y" (equivalently, the choice of the new mean and standard deviation) usually needs some justification.

**Table 5. Scaling formula**

---


$$M' = x + (M * y)$$

$M'$  = new mark

$M$  = original mark

$x$  = origin shift

$y$  = dispersion adjustment

---

## Item banks

The cost-effectiveness of MCQ assessments is only secured by the construction and maintenance of item banks. This follows from the fact the first deployment of a MCQ test is usually costly, and may well be more costly than a conventional essay-type assessment. Subsequently, a required assessment is constructed by sampling most of the desired items from the item bank.

Every assessment must also have a proportion of new items. The "good" new items are placed into the bank, the "bad" ones discarded or re-written and piloted later.

An item bank must be periodically and systematically reviewed and maintained. Out of date items and ones with lower point biserials are discarded. Every time a bank item is used, its performance in a particular assessment must be recorded (the item's difficulty level, its distractor histogram, and its point biserial) and a link established to the test in which it was used.

Separately, the characteristics of each test must be recorded: at least, the number of students, the number of items used in the assessment, any correction for guessing, whether students were instructed to attempt every item, and the assessment's KR-20.

## Example MCQ items

Table 6 illustrates a selection of MCQ templates or fragments (sampled from Bloom, Hastings & Madaus, 1971) designed to address one or other competence drawn from Bloom's cognitive domain. These are illustrated simply to provide ideas for developers who might otherwise wonder how MCQs can assess higher cognitive skills.

**Table 6. Example MCQ templates for Bloom's cognitive domain**

Domain	Scenario or case	Distractors	Possible questions
<b>Comp-rehension</b>	Show a diagram, picture, scenario, vignette, or statements.	A True B Probably true C Insufficient information	Q1 The [X] shows [Y]. Q2 There are N [F]s. Q3 Most of the [Z]s are [W]s.
	For each of the following questions, use the above information to make a decision and indicate your answer using a single letter as follows:	D Probably false E False	
<b>Application</b>	The following are some basic procedures for [X]ing:	A Procedure [1] B Procedure [2] C Procedure [3] D Procedure [4] E Procedure [5]	For each of the following statements, indicate the procedure involved or the most useful procedure to involve. Q1 An [X] which includes a [Y]. Q2 The [X] uses the tool [Z]. Q3 Situation [W] applies while [X]ing. Q4 The goal of the [X] is [G].

Domain	Scenario or case	Distractors	Possible questions
<b>Analysis</b>	Text, diagram etc illustrates a situation or phenomenon [X].  For each question, indicate how the statement illustrates:	A The fundamental purpose of [X] B Achievements incidental to the fundamental purpose of [X] C Evidence of how [X] is regarded by [Y] D Evidence of the involvement of [Z] in [X] E None of the above	Q1 [M] assured [N]. Q2 [X] became associated with [U]. Q3 [P] did much to improve [R]. Q4 [T] was organised around [F]. Q5 [J] demonstrated [K].
<b>Synthesis</b>	The facts [X], [Y], [Z] etc have been found to be the case.  Some procedures to follow include:	A Procedure [R] B Action [S] C Procedure [T] D Action [U] E Procedure [V]	For each statement, indicate the appropriate procedure or action:  Q1 [G] is required. Q2 Because of [H], [J] is to be avoided. Q3 [K] is required as well as a good [L].
<b>Evaluation</b>	Text, diagram, etc outlines a situation or scenario [X].  Views on [X] could be:	A It is thoroughly improper B Proper as regards [Q] but not as regards [W] C Proper as regards [W] but not as regards [Q] D It is thoroughly proper E None of these views clearly applies	In discussing [X], various speakers made the following statements. For each statement, indicate the view that the speaker is likely to hold.  Q1 Hobbs said [V]. Q2 Plato noted [H]. Q3 Mill was right to say [G].

## MCQ test management

An MCQ test is the most popular method of computer-scorable assessment. If an MCQ test, or indeed any form of computer-scorable assessment, is a relatively unfamiliar method to either the developer, lecturer, or student, there are some guidelines which may help, outlined in Table 7.

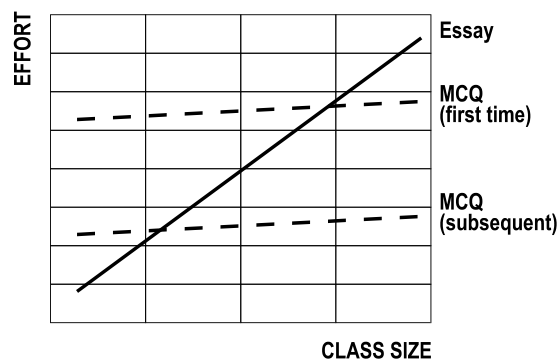
Table 7. Guidelines for the use of MCQ assessment

Stage	Guideline
<b>Preparation</b>	Prepare, distribute, and run mock assessments.  Provide coaching to students on how to answer MCQ tests.
<b>Review</b>	Conduct wide inspections and reviews of MCQ items.
<b>Pilot</b>	Restrict first use of new MCQ items to pilot tests.

MCQ tests require a different distribution of effort in their development, delivery, and marking from essay-type assessments. Table 8 summarises these differences, and Figure 1 illustrates the trade-off of effort against class size.

**Table 8. Effort required for MCQ assessment**

<b>Effort distribution</b>	<p>Effort is front-loaded, put into the setting of the assessment, not the marking.</p> <p>Exact reverse of essay-type assessments.</p> <p>Effort for MCQ tests is largely independent of class size.</p>
<b>Total effort</b>	<p>Total effort most likely to be higher for MCQ than for essay-type assessments in their first application, depending upon class size and number of items required.</p> <p>Subsequent use MCQs take up to much less effort.</p> <p>MCQ tests unlikely to be cost-effective for small classes.</p>



**Figure 1. MCQ vs essay test effort against class size [mcqgraph]**

The following appendices may be useful for MCQ test construction and management:

- Appendix A: National Union of Students' Principles of Effective Assessment.
- Appendix B: Example MCQ.
- Appendix C: Worked example of a MCQ.
- Appendix D: Accessibility and e-assessment.

## **References**

Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. McGraw-Hill.

EASiHE (2010). *EASiHE: e-Assessment in Higher Education*. Gary Wills, Lester Gilbert, Bill Warburton, & David Bacigalupo (Eds). Available at <http://easihe.ecs.soton.ac.uk/>.

Gilbert, L. & Gale, V. (2008). *Principles of E-learning Systems Engineering*. Chandos.

REAQ (2009). *Report on Summative E-Assessment Quality (REAQ)*. Lester Gilbert, Veronica Gale, Bill Warburton, & Gary Wills (Eds). Available at <https://www.webarchive.org.uk/wayback/archive/20140614094217/http://www.jisc.ac.uk/whatwedo/projects/reaq.aspx>

## Appendices

### **Appendix A** ***National Union of Students' Principles of Effective Assessment<sup>1</sup>***

Principle	Comment
Should be for learning, not simply of learning	This positions assessment at the heart of learning rather than it serving as a simple add-on at the end of the process.
Should be reliable, valid, fair and consistent	It is crucial for staff, students and employers to have confidence in the assessment processes and their outcomes.
Should consist of effective and constructive feedback	Effective feedback on assessment is a crucial aspect of assessment processes and a key feature of enhancing the learning process.
Should be innovative and have the capacity to inspire and motivate	Formative assessment practices have the potential to inspire and motivate, and this aspect can be captured by innovative approaches, including those making use of new technology
Should measure understanding and application, rather than technique and memory	Assessments need to have a holistic approach that transcends the particular method being used; only this will truly test and reflect levels of learning.
Should be conducted throughout the course, rather than being positioned as a final event	Positioning assessment as an integral part of the course helps facilitate continuous learning.
Should develop key skills such as peer and reflective assessment	Not only do such mechanisms allow students to receive extra feedback on work beyond that of their tutor, they also help develop the key skill of self-reflection.
Should be central to staff development and teaching strategies, and frequently reviewed	Assessment processes must be innovative and responsive to learners' needs, and as such they need to be central to staff development and teaching strategies.
Should be of a manageable amount for both tutors and students	While assessment should be place in a central role in learning, for it to be effective neither tutor nor student should be overburdened.
Should encourage dialogue between students and their tutors and students and their peers	It is important that students and staff share the same definitions and ideas around standards. This can be fostered by increased dialogue and engagement.

<sup>1</sup> Times Higher Education, 29<sup>th</sup> January 2009. Also see: <https://www.nus.org.uk/en/advice/course-reps/feedback-what-you-can-expect/>

## Appendix B

### Example MCQ

The example addresses the intended learning outcome for a survival skills course:

*At the end of the course the student will be able to analyse an emergency situation, evaluate alternative courses of action, and make a decision that will ensure the survival of all involved for the longest possible time.*

### Structure of an e-assessment question

In the hints and tips section we use the following terms to refer to the different parts of an e-assessment question:

- **Context**  
This provides the information learners need to place the question in context and work out the correct answer. For example, the context may be a case study, a scenario, a diagram, an equation or a role brief.
- **Stem**  
The sentence that poses the question e.g., 'Which one of the following actions is most likely to ensure you all survive?'
- **Instructions**  
How the student should answer the question e.g., 'Click the option you think is correct. Then press Enter.' Instructions also mean the information given to the learner at the start of the e-assessment such as how many questions to answer and any time limit that applies.
- **Correct answer**  
The option or options that demonstrate the learner has mastered the competence being tested by meeting the competence, sometimes called the 'key'.
- **Distractors**  
Incorrect options that the learner needs to consider and discard when working out the correct answer.
- **Feedback**  
The response prepared by the tutor that tells the learners if their selection was correct or incorrect, and explains why.

### Example

<i>Context</i>	You have bailed out of a light aircraft and parachuted safely to the ground. Before ejecting you sent an SOS message. You are alone in a desert. It is 30 degrees centigrade. There is some scrub but no water is in sight. You have an emergency pack that includes a small bottle of water, salt tablets, a protein bar, a spade, a box of matches, a flashlight and a small but very sharp knife. The sun will set in one hour.
<i>Stem</i>	Which one of the following actions is your first priority to maximise your chances of survival?
<i>Instructions</i>	Type in the letter of your answer then click the Done button.
<i>Distractor</i>	a) Burn your parachute to attract attention.
<i>Distractor</i>	b) Dig a hole for shelter overnight.

*Correct answer*      c) Gather scrub for a fire.

*Distractor*          d) Drink the water.

*Feedback*            *If option a) is selected display:*

Unless you are very close to the airfield where a rescue will be launched, we advise you keep the parachute. A rescue attempt is much more likely in daylight. You can use your parachute to attract attention by laying it flat on the sand and holding it in place with stakes cut from the scrub. Your priority for tonight is to keep warm. Gather scrub for a fire while it is still light.

*If option b) is selected display:*

Once the sun sets the sand will get cold very quickly. Also, it is dangerous to dig a hole in sand because the walls are likely to collapse and bury you. Once the sun sets the temperature will drop very quickly. Your priority for tonight is to keep warm. Gather scrub for a fire while it is still light.

*If option c) is selected display:*

Well done, you are correct. Once the sun sets the temperature will drop very quickly. So your priority for tonight is to keep warm. Gathering scrub for a fire while it is still light is the best course of action.

*If option d) is selected display:*

You will need to ration your water until rescue arrives so don't drink it now. Once the sun sets the temperature will drop very quickly. Your priority for tonight is to keep warm. Gather scrub for a fire while it is still light.

## **Appendix C**

### **Worked example of a MCQ**

The following is a poor multiple-choice question. It is followed by an improved version of the same question.

---

### **Intended learning outcome**

The question is intended to address this competence:

*At the end of the course the student will be able to analyse an emergency situation, evaluate alternative courses of action and make a decision that will ensure the survival of all involved for the longest time.*

### **First attempt at writing the question**

#### **Q1: The situation**

You are the captain of a yacht with three crew, Rudy, Clare and Ali. You are sailing around the Azores when a storm blows you off course. You don't know where you are and you have not seen land or another vessel for three days. The yacht is holed by debris at 2300 hours and starts to sink rapidly. You launch the inflatable lifeboat and all the crew board it. The lifeboat is equipped with 3 flares, 1 litre of water, 2 energy bars, a first aid kit, tinder and a box of matches, four life jackets and 2 small paddles. You have 7 minutes to grab supplies before you have to board the lifeboat yourself.

You can only take one of the following. Which one?

- a. Eight tins of food and two litre bottles of water
- b. Four x 4 litre containers of water and one knife
- c. Four space blankets and one fishing rod
- d. One mobile phone and two torches

#### **Feedback seen by students who chose that option**

a	Do you have a can opener? Probably not, so it will be difficult to get into the tins. It is good that you chose some water. The most important consideration is keeping everyone alive for as long as possible. Ideally each person needs 2 litres a day. Option b is better.	1 mark
b	Good, you have maximised your chances of survival.	3 marks
c	I can see why you went for the blankets – it can get chilly at night. And the rod would let you catch fish. But water is the biggest problem. Option b is best.	0 marks
d	A fully-charged mobile would help if you can get reception. Unfortunately you can't. Although the torches are useful at night, they won't help when you get thirsty.	0 marks

#### **Positive aspects of this question**

- Interesting and engaging situation.
- Sense of urgency and an idea of what is at stake.

**Problems with this question**

Some of the problems with this question are:

- Case study description contains information that is not needed now (such as the names of the crew) but does not give key information needed to answer the question correctly (e.g., about mobile phone reception).
- Scenario seems unrealistic, e.g., it is not clear why all the crew left the vessel without taking items such as these and how one person would be able to carry all the items in option a). Further, in seven minutes, a person could load more than one of these combinations, depending on how they were packaged.
- The lifeboat survival kit does not include a torch or a fishing hook. These should be standard items.
- The list of items in the lifeboat could be placed in a pop-up, so there is enough space on the screen to show the question, the student's response and the feedback on one screen.
- No clear instructions about how to answer the question.
- The question does not explain that the choice of items is to maximise the survival of the group.
- The stem specifies 'one of the following' yet all the answer options have two or more items.
- Water appears in two options. This gives an unintentional clue that water is the most important item.
- The learning point that water is the most important item for survival is so central to a survival course, that it seems unlikely that any student would get it wrong.
- The case description does not specify important information such as that there is no land in sight now, or you did not have time to send out a distress call.
- The text for the options is smaller than the case study description, yet is of as much if not more importance.
- The information in the response options is insufficient e.g., what type of knife is it?
- The question assumes the student knows information about the Azores e.g., that it gets relatively cold at night.

The feedback:

- Feedback a) the query about the can opener can be read as sarcastic.
- Feedback b) does not explain how the student has maximised their chances of survival.
- Feedback c) does not explain why option b is best.
- Feedback d) gives further information that was needed in the case description. It does not explain what is the best option and why.

The marks:

- Seem unfair. Although the feedback for option c) gives praise for choosing blankets, no marks are awarded.

## ***Improved version of the question***

### ***Introduction screen***

The following assessment is based on a real scenario. You will be asked five questions. You need to answer every question.

Each question gives you details about the case study and asks you to make a decision. You can get further information by clicking any words that are underlined like this: Far Islands.

Once you have answered, you will get feedback on your choice. This feedback has been prepared by a survival expert.

Every answer that is correct, fully or partially, will earn you marks. The maximum score is 15 marks. These marks do not count towards your final grade for the course. The purpose of these questions is to help you check what you have learned so far.

Click Next when you are ready to start.

[On next screen]

### ***Question 1***

You are the captain of a yacht with two crew members. You are sailing around the Far Islands when a storm blows you off course. Your radio is broken and you have no mobile phone reception. You cannot see land or another vessel. The yacht is holed by debris at 2300 hours and starts to sink rapidly. You are all wearing rain gear and life jackets. You launch the inflatable lifeboat and all the crew board it. The lifeboat has a survival kit.

Type the letters of the following options in the order you plan to grab them (most important first). Your goal is to keep everyone alive as long as possible.

To change your answer, overtype your current response. When you have finished, click Done.

- a. Compass
- b. Cooking pan
- c. Mora knife
- d. Pack of six food tins
- e. Six litre barrel of water
- f. Two blankets

### ***Pop-up for Far Islands***

The Far Islands are a group of islands in the Atlantic Ocean. Many are small and uninhabited. Temperatures at this time of year: high 20.4°C, low 10°C. Average rain fall for month: 3.84 inches.

### ***Pop-up for survival kit***

3 flares, 2 litres of water, 2 energy bars, a first aid kit, tinder and a box of matches in a water-proof bag, 2 small paddles, 1 wind-up torch, 2 fishing hooks, utility knife: (includes two sharp blades, tin opener, spoon and screw driver), billy can.

**Feedback**

edcfab	<p>Well done, you have chosen the order that our expert advises will maximize your chances of survival.</p> <ol style="list-style-type: none"> <li>1. Water: ideally every person needs 2 litres a day.</li> <li>2. Food: important if you don't find land soon.</li> <li>3. Mora knife: strong enough to gut fish, cut wood for and defend you from animals.</li> <li>4. Blankets: it gets chilly at night. The body needs more water if it gets cold.</li> <li>5. Compass: if you decide to row, you'll need the compass to head in one direction.</li> <li>6. Cooking pan: this could be used to collect condensation. As you have a billy can it is not a priority.</li> </ol> <p>See chapter 2 of 'Survival Now' for more details.</p>	3 marks
e appears in first or second position	<p>Good, you placed water high in your priorities. This is the order our expert recommends and why:</p> <ol style="list-style-type: none"> <li>1. Water: ideally every person needs 2 litres a day.</li> <li>2. Food: important if you don't find land soon.</li> <li>3. Mora knife: strong enough to gut fish, cut wood for and defend you from animals.</li> <li>4. Blankets: it gets chilly at night. The body needs more water if it gets cold.</li> <li>5. Compass: if you decide to row, you'll need the compass to head in one direction.</li> <li>6. Cooking pan: this could be used to collect condensation. As you have a billy can it is not a priority.</li> </ol> <p>See chapter 2 of 'Survival Now' for more details.</p>	1 mark
Any other order	<p>Our expert recommends water is placed first. Compare your order and rationale to the order he recommends:</p> <ol style="list-style-type: none"> <li>1. Water: ideally every person needs 2 litres a day.</li> <li>2. Food: important if you don't find land soon.</li> <li>3. Mora knife: strong enough to gut fish, cut wood for and defend you from animals.</li> <li>4. Blankets: it gets chilly at night. The body needs more water if it gets cold.</li> <li>5. Compass: if you decide to row, you'll need the compass to head in one direction.</li> <li>6. Cooking pan: this could be used to collect condensation. As you have a billy can it is not a priority.</li> </ol> <p>See chapter 2 of 'Survival Now' for more details.</p>	0 marks

## ***Appendix D***

### ***Accessibility and e-assessment***

This section was first published as part of the Report on E-Assessment Quality (REAQ) and is reproduced here with permission.

Most students, once they have become familiar with online testing prefer this medium. However, in the past guidance about making sure computer based or online assessments are accessible to disabled students often avoided the technological issues that could arise and concentrated on the inclusion of extra time for completion or the provision of alternative formats such as paper based examinations. There are situations where these suggestions may not always prove to be helpful and it is vital to be aware of individual preferences which have been highlighted in the check list below.

### ***Key Points to Look for in Online Assessment***

Keyboard navigation relates to the use of Internet Explorer. Other browsers may differ.

---

<b>Criterion</b>	Is text re-sizeable?
<b>Why it Matters</b>	People with poor sight may need larger text sizes. People with tunnel vision may need smaller text sizes.
<b>How to check it</b>	Using Internet Explorer, select View > Text size and select either larger or smaller text size

---

<b>Criterion</b>	Can the active elements of the quiz (e.g. selection buttons) be selected using the keyboard only?
<b>Why it Matters</b>	Some users find difficulty using a mouse.
<b>How to check it</b>	Click on the Tab key. Successive clicks should highlight the different elements of the page in turn. When an option has been highlighted, pressing the Return key should activate the selection.

---

<b>Criterion</b>	Do the active elements come up in a logical sequence when selected using the keyboard only?
<b>Why it Matters</b>	Users relying on screen readers use the Tab key to navigate through options. If the options present in a confusing sequence it can be difficult to make sense of the assessment
<b>How to check it</b>	Click on the Tab key. Successive clicks should highlight the different elements of the page in turn. See if these come up in logical order from top to bottom and left to right.

---

<b>Criterion</b>	Do selectable options (e.g. radio buttons or check boxes) appear before or after the descriptive text?
<b>Why it Matters</b>	Users relying on screen readers want to select an option AFTER they have had it read to them, not before.
<b>How to check it</b>	Inspect the page layout to see if the selection buttons are before or after the text.

---

<b>Criterion</b>	Where images are used, do they have appropriate descriptions that inform the user without giving away the answers?
<b>Why it Matters</b>	Users relying on screen readers may not be able to see the image so a text description is required. However, the description must not give away the answer - for example the question "Which of the following images represents butane?" should not have the description "Butane molecule" as an [Alt] tag on the image but rather a description of the number and nature of atomic bonds
<b>How to check it</b>	Hover the mouse cursor over the image and see if a screen tip description comes up.

---

This has been taken from JISC TechDis: <https://www.jisc.ac.uk/guides/open-educational-resources/accessibility-considerations>