Whither e-assessment in the mathematical sciences: a critical view from the edge

Sally Jordan   @SallyJordan9

EAMS September 2016

The Open University

Life-changing Learning

A view from the edge?

- I am not a mathematician.
- I am not a technical expert.

- I am passionate about students and learning.
- I have used online computer-marked assessment with computer-generated feedback in my teaching since 2002 (initially on *Maths for Science* and subsequently on a range of other modules).
- From 2006, I evaluated the use of automatically marked questions in which students give their answer as a free-text phrase or sentence, using a range of software. This led to the Moodle "Pattern Match" question type.

My context: UK Open University

- Founded in 1969
- Supported distance learning
- 200 000 students, mostly studying part-time
- Undergraduate modules are completely open entry, so students have a wide range of previous qualifications
- Normal age range from 18 to ??
- 20 000 of our students have declared a disability of some sort
- 13 000 of our students live outside the UK

iCMA = interactive computer-marked assignment

TMA = tutor-marked assignment

My plan

❑Are we delivering high quality e-assessment? What can we do to improve things?

❑More about Pattern Match.

❑What does the future hold?

My plan

❖ What do we mean by high quality e-assessment?

❖ What is (e) assessment for?

❑ Are we delivering high quality e-assessment? What can we do to improve things?

❑ More about Pattern Match.

❑ What does the future hold?

My plan

❖ What do we mean by high quality e-assessment?

❖ What is (e) assessment for?

- *What have other keynote speakers said?*

- *What do the experts say?*

- *What do our students say?*

- *What do you say?*

❑ Are we delivering high quality e-assessment? What can we do to improve things?

❑ More about Pattern Match.

❑ What does the future hold?

To get you thinking…

"Speed talking" [idea courtesy of Ian Bearden]

Find yourself a partner, and decide which of you is Person A and which is Person B.

Be prepared to talk for 20-30 seconds on a topic…

…when the slide changes.

Person A

E-assessment

Person B

Assessment for Learning

Person A

Learning analytics

Person B

High quality e-assessment

STOP!

What do the experts say?

Assessment can define a "hidden curriculum" (Snyder, 1971).

Whilst students may be able to escape the effects of poor teaching, they cannot escape the effects of poor assessment. (Boud, 1995).

Summative assessment is itself "formative". It cannot help but be formative. This is not an issue. At issue is whether that formative potential of summative assessment is lethal or emancipatory. Does summative assessment exert its power to disrupt and control, a power so possibly lethal that the student may be wounded for life? (Barnett, 2007).

What have our other keynote speakers said?

Michael: "Ask the questions you should, not just the ones you can."

Christian: "The experience of *using* e-assessment…is ignored at your peril."

Chris: "Where are the limits of automatic assessment in the future?"

What do our students say?

Comments from students

- *I discovered, through finding an error in the question, that not everybody was given the same questions. I thought this was **really unfair** especially as they failed to mention it at any point throughout the course.*

- *I find them **petty** in what they want as an answer. For example, I had a question that I technically got numerically right with the correct units only I was putting the incorrect size of the letter. So I should have put a capitol K instead of a lower case k or vice versa, whichever way round it was. Everything was correct except this issue.*

Thankfully, these students were happy with computer-marked assessment in general, but particular questions had put them off.

Comments from students

- *A brilliant tool in building confidence*
- *It's more like having an online tutorial than taking a test*
- *Fun*
- *It felt as good as if I had won the lottery*
- *Not walkovers, not like an American-kind of multiple-choice where you just go in and you have a vague idea but you know from the context which is right*

And from a tutor

- *Even though each iCMA is worth very little towards the course grade my students take them just as seriously as the TMAs. This is a great example of how online assessment can aid learning.*

"When we consider the introduction of e-assessment we should be aware that we are dealing with a very sharp sword" (Ridgway, 2004).
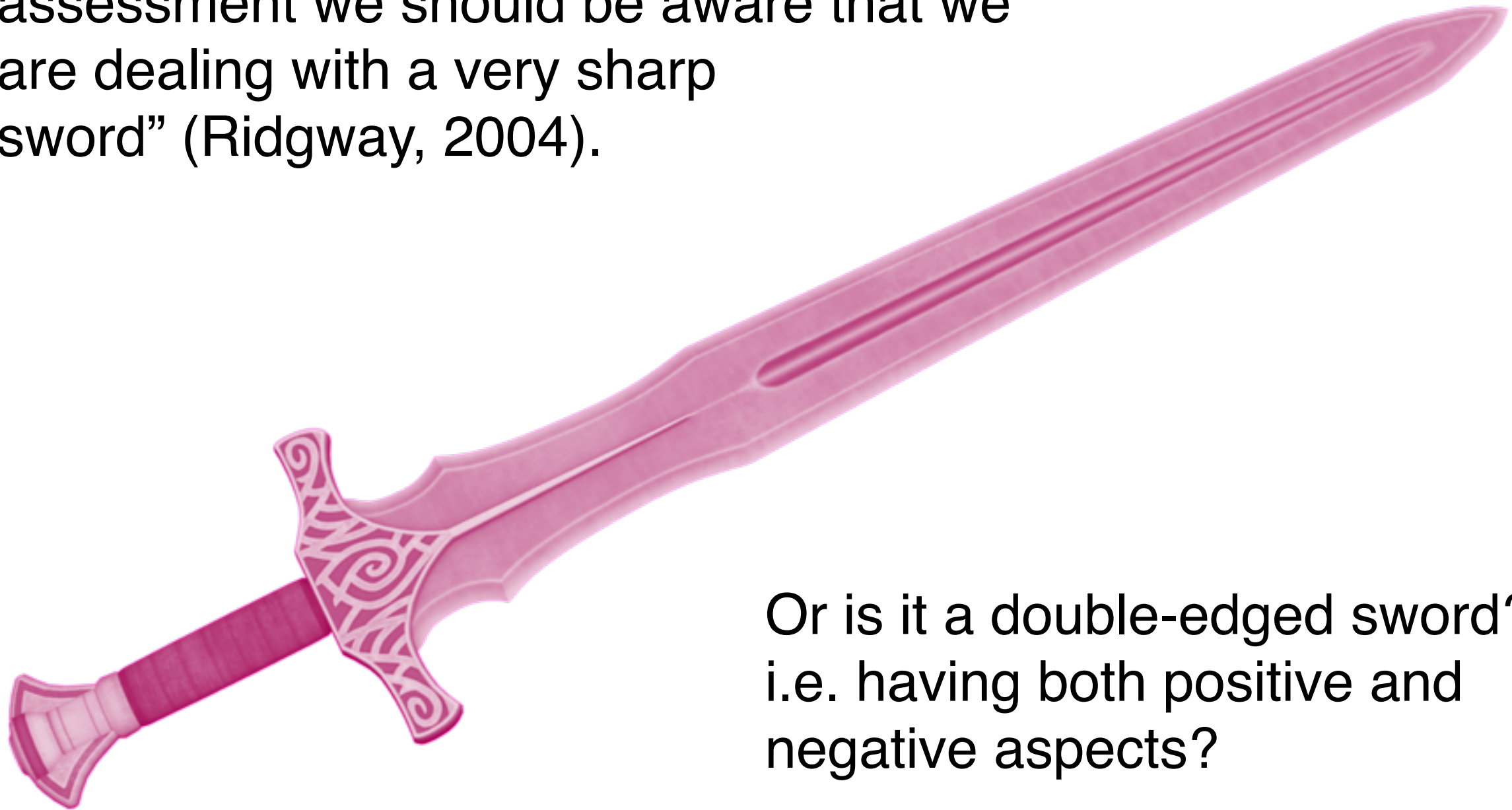
Or is it a double-edged sword? i.e. having both positive and negative aspects?

To maximise the positive…

- Make your e-assessment both efficient and effective. "*Efficiency is doing this right; effectiveness is doing the right things*." Peter Drucker

- Don't be limited in your ideas.

- But don't be beguiled by a wish to use the latest technology. "*Students First.*" Open University strategy.

So, what is e-assessment?

Definition can include any use of a computer as part of any assessment-related activity (JISC, 2006).

So includes:

- "Electronic management of assessment"
- Audio/video feedback
- ePortfolios
- Use of blogs or wikis in assessment
- Assessment of online forums
- Use of computers for exams
- Interactive online computer-marked assessment with computer-generated feedback

Not all computer-marked assessment is the same

To improve quality:

- Think about why you want to use computer-marked assessment. *Assessment of Learning* or *Assessment for Learning*?
- Think about your assessment design; how will you integrate it?
- Use appropriate question types
- Write better questions with better feedback
- Use an iterative design process

Potential advantages of computer-marked assessment

- To save staff time
- To save money
- For constructive alignment with online teaching
- To make marking more consistent ('objective')
- To enable feedback to be given quickly to students
- To provide students with extra opportunities to practise
- To motivate students and to help them to pace their learning
- To diagnose student misunderstandings

Potential disadvantages of computer-marked assessment

- May encourage a surface approach to learning
- May not be authentic
- There is no tutor to interpret the student's answer and to deliver personalised feedback

- Tends to mark "an answer" rather than the working
- Issues with symbolic notation for mathematics and related disciplies

Why have I used computer-marked assessment?

- In my work, the focus has been on 'assessment for learning', so feedback and giving students a second and third attempt is important (Gibbs & Simpson, 2004-5).

- We aim to 'provide a tutor at the student's elbow' (Ross et al., 2006).

- However, a summative interactive computer-marked assignment that ran for the first time in 2002 is still in use, and has been used by around 16,000 students.

Assessment design

- From Twitter yesterday:

In two sessions on #flipping #EAMS2016. Really pleased that the conference is about more than question design.

- Good question design is a necessary but not sufficient condition for good e-assessment.

Use appropriate question types

- Multiple-choice

- Multiple-response

- Drag and drop

- Matching

- True/false

- Hotspot

- Free text: for numbers, letters, words, sentences

Note: You need to think about what your e-assessment system supports.

My work with short-answer free-text questions

- Had the original goal of extending the types of computer-marked assessment that was available;

- Focused on 'Assessment for Learning' i.e. feedback to students and an opportunity to have another go;

- Developed answer-matching **using responses from hundreds and thousands of real students**;

- Used two different software approaches;

- Both worked surprisingly well; ideas now incorporated into Moodle Pattern Match.

Pattern Match is an algorithmically based system

- so a rule might be something like

Accept answers that include the words 'high', 'pressure' and 'temperature' or synonyms, separated by no more than three words

- This is expressed as:

else if ((m.match("mowp3", "high|higher|extreme|inc&|
immense_press&|compres&|[deep_burial]_temp&|heat&|
[hundred|100_degrees]")

matchMark = 1; whichMatch = 9;

- 10 rules of this type match 99.9% of student responses

Example of a short-answer question

If the distance between two electrically charged particles is doubled, what happens to the electric force between them? Be as specific as possible.

*Please give your answer as a **short** phrase or sentence.*

The force will decrease.

Enter answer

Your answer appears to be incorrect or incomplete in some way.

You are on the right lines. You are correct to say that the strength of the force decreases, but by how much? Coulomb's Law states that the electric force between two charged particles is inversely proportional to the square of their separation (see Book 7 Section 10.1). So when the distance between the particles is doubled, what happens to the electric force between them?

Try again

Example of a short-answer question *cont.*

If the distance between two electrically charged particles is doubled, what happens to the electric force between them? Be as specific as possible.

*Please give your answer as a **short** phrase or sentence.*

The force will halve.

Enter answer

Your answer still appears to be incorrect or incomplete in some way.

You are correct to say that the strength of the force decreases, but not to say that it halves. Coulomb's Law states that the electric force between two charged particles is inversely proportional to the square of their separation (see Book 7 Section 10.1). So when the distance between the particles is doubled, what happens to the electric force between them?

Try again

Example of a short-answer question *cont.*

If the distance between two electrically charged particles is doubled, what happens to the electric force between them? Be as specific as possible.

Please give your answer as a **short** *phrase or sentence.*

The force will decrease by a factor of four.

Enter answer

Your answer is correct.

Coulomb's Law states that the electric force between two charged particles is inversely proportional to the square of their separation (see Book 7 Section 10.1). So when the distance between the particles is doubled, the electric force between them is reduced by a factor of four i.e. it is a quarter of its original value.

Next question

Simple but not that simple?

A ball is thrown vertically upwards into the air. What is the main energy conversion that takes place as the ball rises towards its maximum height?

Note: you should only consider the energy conversions that take place after the ball has been thrown and before it reaches its maximum height.

*Please give your answer as a **short** phrase or sentence.*

Kinetic energy is converted to gravitational potential energy.

Enter answer

Your answer is correct.

The main energy conversion as the ball rises is from kinetic energy (energy of motion) to gravitational potential energy. Gravitational energy is a type of potential (stored) energy and it will have a maximum value when the ball is at its maximum height. At the same point the ball will be instantaneously stationary, so it will have no kinetic energy.

Energy conversions are discussed in Book 3 Chapter 2.

Next question

Simple but not that simple?

A ball is thrown vertically upwards into the air. What is the main energy conversion that takes place as the ball rises towards its maximum height?

Note: you should only consider the energy conversions that take place after the ball has been thrown and before it reaches its maximum height.

*Please give your answer as a **short** phrase or sentence.*

> Gravitational potential energy is converted to kinetic energy.

Enter answer

Your answer appears to be incorrect or incomplete in some way.

You are correct to say that the two types of energy involved are gravitational potential energy and kinetic energy, but you need to describe the energy conversion more precisely. Energy conversions are discussed in Book 3 Chapter 2.

Try again

Simple but not that simple?

A raindrop falls vertically with a constant speed. What does this tell you about the forces acting on the raindrop?
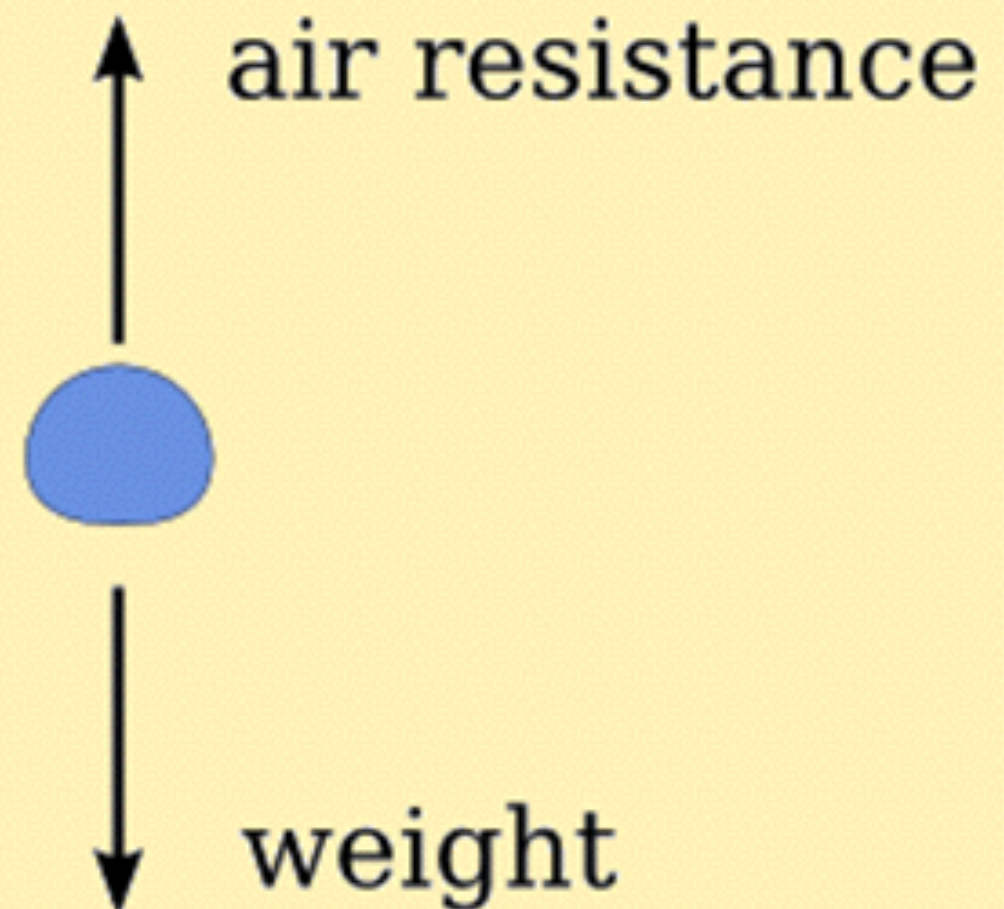
*Please give your answer as a **short** phrase or sentence.*

There are no unbalanced forces acting.

Enter answer

Your answer is correct.

Since the raindrop is falling with constant velocity the forces on it must be balanced. This is a consequence of Newton's First Law of Motion.

↑ air resistance

↓ weight

Simple but not that simple?

A raindrop falls vertically with a constant speed. What does this tell you about the forces acting on the raindrop?

*Please give your answer as a **short** phrase or sentence.*

The forces are unbalanced.

Enter answer

Your answer still appears to be incorrect or incomplete in some way.

What does Newton's First Law of Motion tell you about the forces acting on an object moving with a constant velocity?

See Block 3 Section 4.1.

Try again

Question types in use (2012)

TOP TEN MOODLE QUESTION TYPES (Worldwide)	Number	%
Multiple choice	40,177,547	74.85
True/false	6,462,669	12.04
Short-answer	3,379,336	6.30
Essay	2,321,918	4.33
Matching	551,404	1.03
Multi-answer	341,988	0.64
Description	149,303	0.28
Numerical	138,761	0.26
Calculated	103,103	0.19
Drag-and-drop matching	26,117	0.05
TOTAL	53,675,508	100

Hunt, T. (2012). Computer-marked assessment in Moodle: Past, present and future. Paper presented at the International CAA Conference, Southampton, July 2012.

Questions attempted at OU, 01-01-2015 to 08-03-2016

Question type	# Qs attempted	Percentage
multichoice	2391427	34.22%
stack	1077096	15.41%
oumultiresponse	569182	8.14%
ddwtos	562500	8.05%
numerical	544174	7.79%
description	488968	7.00%
match	412094	5.90%
shortanswer	255738	3.66%
truefalse	201979	2.89%
gapselect	192300	2.75%
opaque	87475	1.25%
combined	68180	0.98%
pmatch	7214	0.10%

Constructed response or selected response?

- The most serious problem with selected response questions is their lack of authenticity: "Patients do not present with five choices" (Mitchell et al., 2003) quoting Veloski (1999).

- But even relatively simple selected response questions can lead to "moments of contingency" (Black & Wiliam, 2009) enabling "catalytic assessment", the use of simple questions to trigger deep learning (Draper, 2009)

But be careful with question wording…

What's the answer?

The bfeld links to the mnoge by means of a

A   elland

B   angaster

C   tanag

D   introdoll

E   ussop

Be careful with question wording

A snow flake falls vertically with constant speed. What does this tell you about the forces acting on the snow flake?

The forces acting on the snow flake are gravity and air resistance and they are ballanced.

You should give your answer as a short phrase or sentence. Answers of more than 20 words will not be accepted.

Enter answer

Your answer is correct.

Since the snow flake is falling with constant velocity the forces on it must be balanced. This is a consequence of Newton's first law of motion. See Book 2 Section 14.1.1.

↑ air resistance

weight
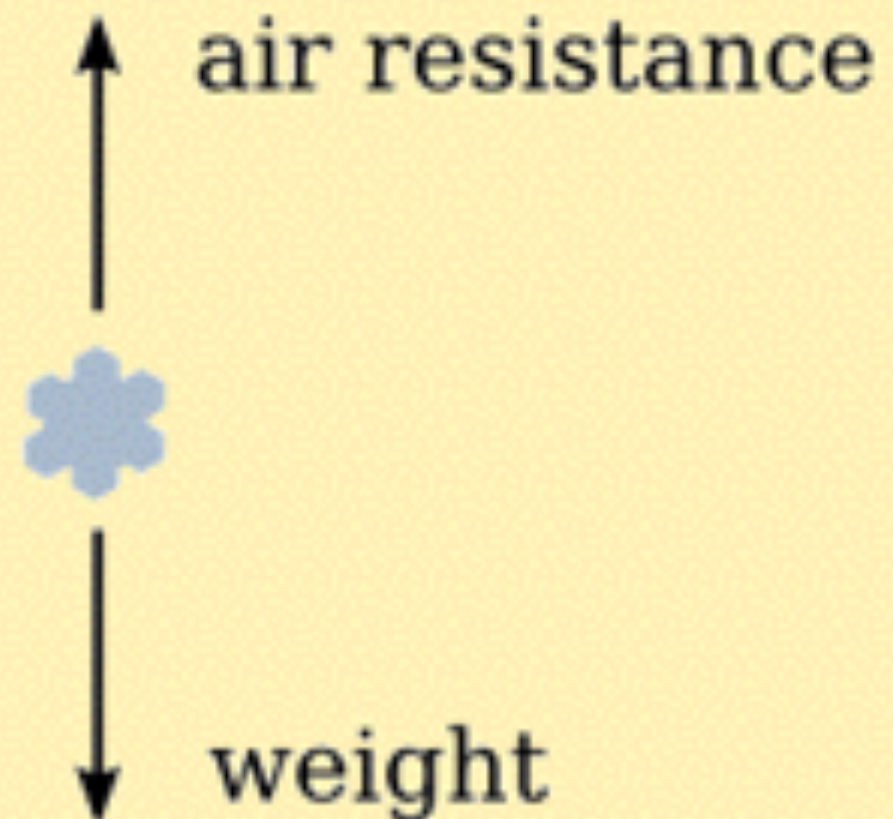
Next question

Our advice to question authors

- Think about how you want your assessment to be embedded within the module
- Think about what question type to use (selected response or constructed response)
- Make sure that your question is carefully worded
- Think about your feedback
- Think about providing variants of the questions
- Check your questions
- Get someone else to check your questions
- Modify your questions in the light of student behaviour the first time they are used.

Monitor question performance
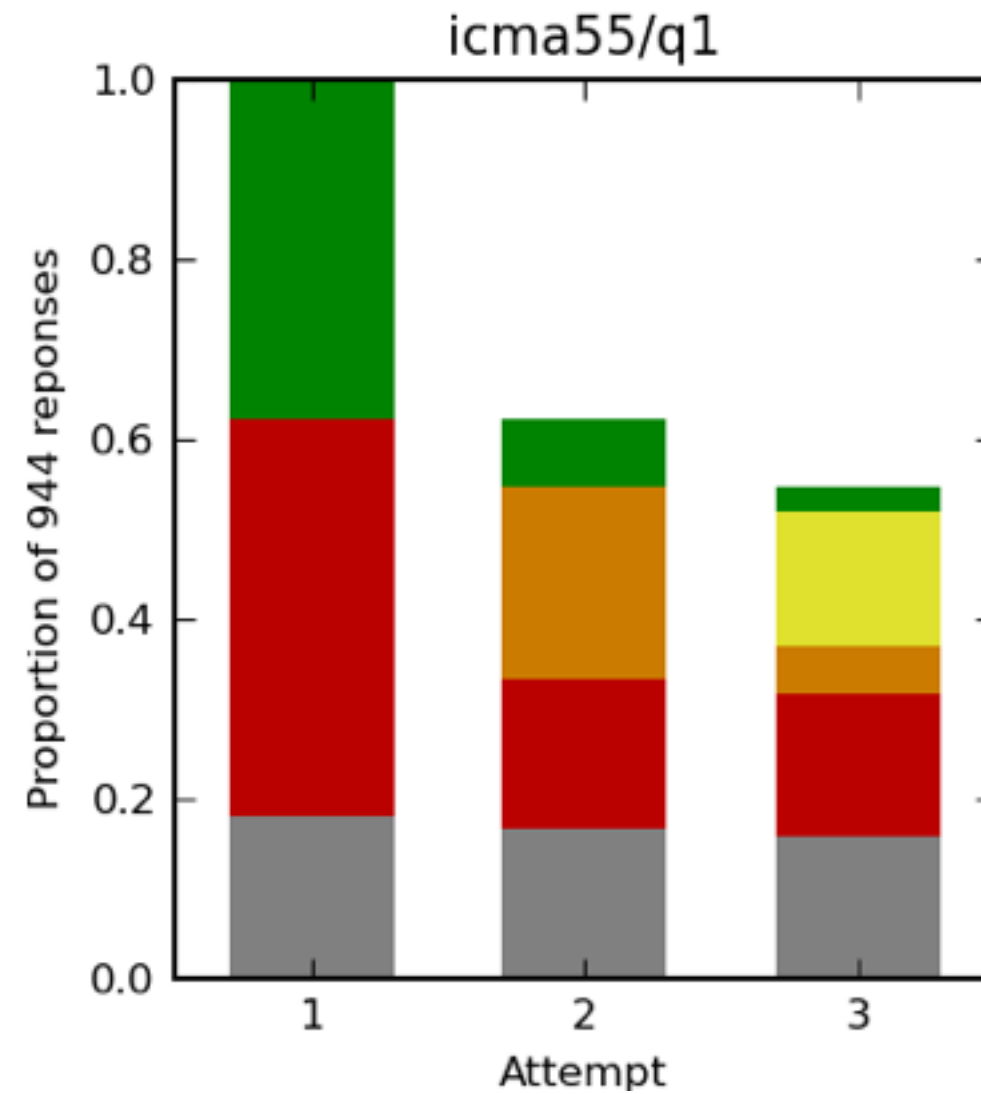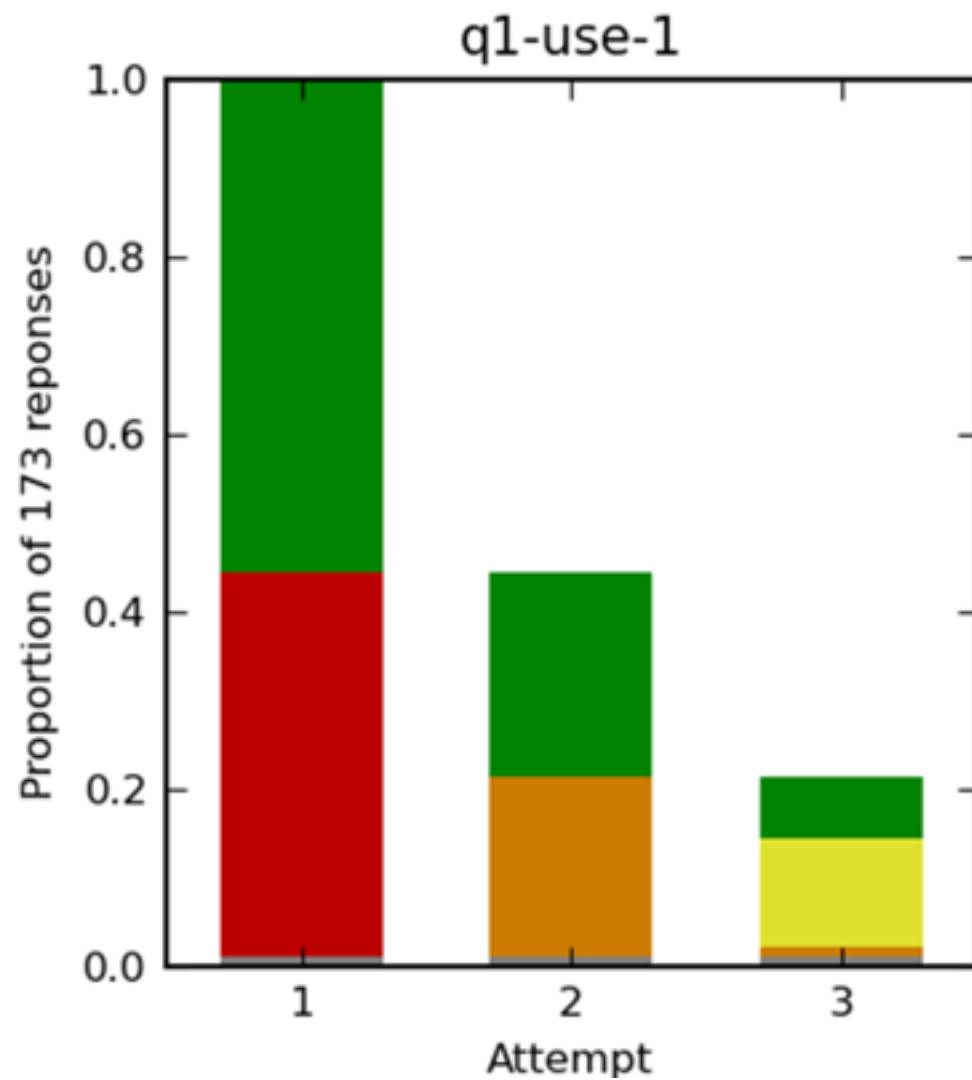
Monitor question performance

What is $\frac{2}{3} + \frac{5}{7}$ expressed as a single fraction? You should give your answer in the simplest possible form.

$$\frac{2}{3} + \frac{5}{7} = \frac{29}{42}$$

Enter answer

Your answer is still incorrect.

$$\frac{2}{3} + \frac{5}{7} = \frac{2 \times 7}{3 \times 7} + \frac{5 \times 3}{3 \times 7}$$

$$= \frac{14 + 15}{3 \times 7}$$

$$= \frac{29}{21}$$

Addition of fractions is discussed in *Maths for Science* Section 1.2.2.

Next question

So we have done quite well

- …helping students directly to improve their understanding and learning more about their misunderstandings

- However writing good questions takes a lot of time and therefore money

Two possible solutions:

- Use machine-learning to develop the answer matching (especially for short-answer free-text questions)
- Share questions

Pattern Match: recent developments

To assist with the authoring of Pattern Match questions, the following have been added:

- A rule creation assistant
- Semi-automated authoring of rules

As part of research into marking student responses to short answer questions, Alistair Willis developed the Amati system which supports question authors in the development of 'rules' for automatic marking (Willis, 2015). This has now been incorporated into Moodle Pattern Match.

Why don't we collaborate more?

"Sharing questions is one of those things which is easy to say we'd like but turns out to be very difficult in practice."

- Some questions are systems dependent (so need interoperability: Question and Test Interoperability (QTI))
- Questions may be context dependent e.g. refer to other resources, assume particular previous knowledge.

Is a solution to share questions and allow others to edit them for their own use?

Note: questions may be confidential (especially if in high-stakes summative use)

How far is it appropriate to go?

- It is technically possible to get good answer matching for some quite sophisticated question types e.g. essays.

- But Perelman (2008) trained students to obtain good marks for a computer-marked essay by "tricks".

- Computer-marked assessment is not a panacea.

"*If course tutors can be relieved of the drudgery associated with marking relatively short and simple responses, time is freed for them to spend more productively, perhaps in supporting students in the light of misunderstandings highlighted by the e-assessment or in marking questions where the sophistication of human judgement is more appropriate*"(Jordan & Mitchell, 2009).

References

Barnett, R. (2007). Assessment in higher education: An impossible mission? In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education* (pp. 29-40). London: Routledge.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5-31.

Boud, D. (1995). Assessment and learning: Contradictory or complementary? In P. T. Knight (Ed.), *Assessment for learning in higher education* (pp. 35-48). London: Kogan Page.

Draper, S. (2009). Catalytic assessment: Understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology*, 40(2), 285-293.

Gibbs, G. & Simpson, C. (2004-5). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, *1*, 3-31.

Mitchell, T., Aldridge, N., Williamson, W., & Broomhead, P. (2003). Computer based testing of medical knowledge. In Proceedings of the 7th International Computer-Assisted Assessment (CAA) Conference, Loughborough, 8th-9th July 2003. Retrieved from http://caaconference.co.uk/pastConferences/2003/proceedings

Perelman, L. (2008). Information illiteracy and mass market writing assessments. College Composition and Communication, 60(1), 128-141.

Ridgway, J., McCusker, S., & Pead, D. (2004). *Literature review of e-assessment*. Bristol: Futurelab.

Snyder, B.R. (1971). *The hidden curriculum*. New York: Alfred A. Knopf.

References to systems discussed in detail

Butcher, P. G. & Jordan, S. E. (2010). A comparison of human and computer marking of short free-text student responses. *Computers & Education*, 55(2), 489-499.

Hunt, T. J. (2012). Computer-marked assessment in Moodle: Past, present and future. In Proceedings of the 2012 International Computer Assisted Assessment (CAA) Conference, Southampton, 10th-11th July 2012. Retrieved from http://caaconference.co.uk/proceedings/

Jordan, S. E. & Mitchell, T. (2009). E-assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology*, *40*(2), 371-385.

Ross, S. M., Jordan, S. E. & Butcher, P. G. (2006). Online instantaneous and targeted feedback for remote learners. In C. Bryan & K. Clegg (Eds.), Innovative Assessment in Higher Education (pp. 123-131). London: Routledge.

Willis, A (2015). *Using NLP to support scalable assessment of short free text responses* [Online]. Available at http://aclweb.org/anthology/W/W15/W15-0628.pdf (Accessed 5 August 2016).

Much of what I have said is discussed in more detail in:

Jordan, S. E. (2014). *E-assessment for learning? Exploring the potential of computer-marked assessment and computer-generated feedback, from short-answer questions to assessment analytics*. PhD thesis. The Open University. Retrieved from http://oro.open.ac.uk/4111