E-assessment: screen design and question difficulty

Changing the way we see test-items in a computer-based environment

Matt Haigh: mrh47@cam.ac.uk

Lessons with a penguin

James is a newly-qualified teacher who is enthusiastic about using technology in the classroom. He has persuaded the head of department to purchase a very impressive software package for assessing his students' science knowledge. The software is particularly exciting to the animated penguin that helps them navigate through the features of the software and gives them feedback on their answers.

After several lessons he looks at the students' assessment results. Following a promising rise in their scores, he notices a sudden downturn in performance. Several of his star students are performing badly on the assessments. He checks the software and observes his students more closely in the following lesson. He observes that the student's faces are blank and, following a string of incorrect answers, he asks one of the students what is going on.

“Well sir, you see, when you get a question right, the penguin just gives you a little smile, but when you get a question wrong, he rolls off his tail and makes a big splash in the water which is really funny, so I'm just glad of getting the questions wrong so I can see the penguin go wet.”

Stage 2 - Qualitative

Two focus groups (each with four students) were held following the online test.

Student comments were divided into two categories: reasons given for features making an item easier and reasons associated with features that make an item more difficult. The word clouds below are based on a word-frequency analysis where font size directly corresponds to the word-frequency.

Analysis and implications

Measures of difficulty

Significant differences were reported in the measure of item difficulty in any of the ten items modified in the test. This implies that students are able to deconstruct their test knowledge through a variety of screen formats and interactions.

Student perceptions

Student perceptions in many cases strongly articulate perceived differences in the difficulty of items. On some occasions students were able to articulate clear reasons relating to both versions of the item that have an influence on difficulty.

Implications for practice

The qualitative data indicated that students are able to handle a range of interactions for providing evidence of their abilities which is significant for the design of future online tests. However, the qualitative data indicate that some forms are perceived differently by students in terms of their difficulty and that this is worth investigating further.

Further research

Evidence that the best of the focus groups of students (i.e. those with poor ICT skills or of lower ability) may be more susceptible to changes in the qualitative phase of the study. The study was undertaken with a limited number of students, from one specific context. Further research could be undertaken with a broader range of item types and in a wider set of contexts to identify whether the findings were more generalisable.

Parallel test forms – Spot the difference...

Key to results:

- Description of the modification to screen design in the parallel forms of the test
- Indicates whether the difference in question difficulty is significant (p≤0.05)

A subset of five out of the ten modified items in the study are shown below.

Qualitative analysis – Word clouds

Words associated with “Easiness”

“Most people would know that [you can click on the tabs], but not if you are in a rush in your exam...on the last few questions you might not know the words like actually on the page. Instead of just clicking and ruffling down at the choices.”

Words associated with “Difficulty”

“On the other hand, it’s easier because you can see all the words like actually on the page. Instead of just clicking and ruffling down at the choices.”

---

This study was undertaken as part of a Masters programme in educational research at the Faculty of Education, University of Cambridge. A full version of the Masters thesis along with samples of the online-tests are available from the following website:

www.assessmentresearch.org.uk