



CAMBRIDGE ASSESSMENT

An investigation into the impact of screen design on computer-based assessments

Matt Haigh

Research Division, Cambridge Assessment

Haigh.M@cambridgeassessment.org.uk

Paper presented at the British Educational Research Association annual conference, University of London Institute of Education, September 2011.

Background

This research is set in a context of increasing technology use across all educational sectors and a drive to apply technology in the field of educational assessment. The current educational research literature links computer-based assessment with themes of formative assessment, validity, access, efficiency and motivation. This study tackles some of the areas of validity in computer-based testing which currently possess limited research evidence.

Research Questions

The aim of this study was to investigate how aspects of screen design in a computer-based assessment may affect the difficulty of test items and to gain some insight from test users into the nature of any observed effects. The research questions addressed were:

- What are the effects of changing the student-item interaction on measures of item difficulty with a computer-based test item?
- What insights are given by the test users for any potential sources of construct-irrelevant variance in computer-based test items?

Methods

Parallel forms of a computer-based assessment were administered to 112 students from 6 schools. Questions in the second parallel form contained a modification of the screen environment (such as layout or method of student interaction). Difficulty measures of each question were calculated and tested for significant differences. In addition, two focus groups of students were used to build up insights into the users' view of the potential impact of the screen environment on the test experience.

Analytical/Theoretical Framework

The research focuses on the concept of 'construct-irrelevant variance' formulated by Messick (1989) in his work on the validity of assessment outcomes. The researcher adopted a pragmatic approach to the study; using a sequential mixed-methods approach to investigate two perspectives. Firstly, the quantitative experimental approach was used to examine the difficulty measures associated with the test. Secondly, a qualitative approach was used to build up a picture of the users' experiences.

Research Findings

The first strand of the research identified no significant differences in the difficulty of items as a result of modifications to the screen environment. However, the second strand indicated that many test items were perceived differently by students taking the test and provided insights into



how aspects of screen design have an impact on the students' perceptions of the questions. This has implications for educational assessment designers, particularly where the validity of the test outcomes are affected.

References

Messick, S. (1989). Validity. In R. L. Linn, *Educational Measurement* (3rd ed.). New York: Macmillan.