Computers & Education 85 (2015) 35–48

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

What makes an item more difficult? Effects of modality and type of visual information in a computer-based assessment of scientific inquiry abilities

Hsin-Kai Wu [a,*], Che-Yu Kuo [a], Tsung-Hau Jen [b], Ying-Shao Hsu [a]

[a] Graduate Institute of Science Education, National Taiwan Normal University, P.O. Box 97-27, Taipei, 11699, Taiwan, ROC

[b] Center of Science Education, National Taiwan Normal University, 88 Sec. 4, Ting-Chou Road, Taipei, 11677, Taiwan, ROC

A R T I C L E I N F O

Article history:
Received 22 November 2014
Received in revised form
7 January 2015
Accepted 22 January 2015
Available online 18 February 2015

Keywords:
Architectures for educational technology system
Secondary education
Applications in subject areas

A B S T R A C T

Effects of multiple external representations on teaching and learning have been widely researched; however, relatively little is understood about how different types of representations used for item presentation influence students' performances in computer-based assessments, particularly in those evaluating complex abilities. We designed a multimedia-based assessment for secondary school students that focused on scientific inquiry abilities (i.e., questioning, experimenting, analyzing, and explaining). Through the participation of 1218 students (561 8th graders and 657 11th graders), the balanced arrangement of test booklets, and the use of a generalized partial credit Rasch model, this study aimed at investigating how the modality of representations for item presentation (i.e., dynamic and static) and type of visual information (i.e., context and content) conveyed by the representations affected the item difficulty of the multimedia-based assessment. The results showed that overall items became slightly more difficult for students when the item presentation was static. Also, the interaction effect between modality and grade was significant; when item presentation was changed from dynamic to static, the items became easier to the 8th graders, whereas they were more difficult to the 11th graders. The results suggested that older students might have more cognitive resources to retrieve information from dynamic displays to solve the assessment tasks successfully. Additionally, while all interactions between modality and type of information on items with context visuals were not significant, there were significant interactions on four items with content visuals. A further examination of the items suggested that the tasks and the phenomena involved in the items may influence how modality affected the item difficulties when content visuals were used. These results implied that researchers need to pay special attention on the design of content visuals because they appear to be more influential than context ones to students' performances of science assessments.

1. Introduction

The benefits of computer-based assessments (CBAs) have been well documented for decades (e.g., Bunderson, Inouye, & Olsen, 1989; Scheuermann & Björnsson, 2009). In addition to the efficiencies of test production, delivery, and administration, CBAs have great potential for increasing fidelity to the constructs of interest and provide opportunities to measure learning outcomes that are hardly or not possible to be tested in paper-and-pencil assessments (Kuo & Wu, 2013; Organization for Economic Co-operation and Development [OECD], 2010). For example, in science education, fundamental abilities of inquiry, such as asking scientific questions, planning an experiment, and constructing explanations, have been viewed as important goals for science curricula (Ministry of Education [MOE], 1999, 2008; National Research Council [NRC], 2000) but evaluating these complex abilities in science usually require performance assessments and the development and implementation of reliable and valid performance assessments could be challenging and time-consuming (Ruiz-Primo &

* Corresponding author. Tel.: +886 2 77346805; fax: +886 2 29333018.
E-mail address: hkwu@ntnu.edu.tw (H.-K. Wu).

http://dx.doi.org/10.1016/j.compedu.2015.01.007

Shavelson, 1996). In recent years, therefore, more and more assessments of complex competences in science at both national and international levels have taken advantage of computer technology and adopted dynamic and interactive media, including the Programme for International Student Assessment (PISA) in science (OECD, 2010) and National Assessment Educational Progress (NAEP) in the United States (Quellmalz, Timms, Silberglitt, & Buckley, 2012).

However, although computer technology has been assumed to bring benefits to educational assessments, the potential is not fully realized and explored. As found in their review of simulations in science, Scalise et al. (2011) indicated that "few or no assessment approaches described in the research articles took advantage of the simulation or virtual context for assessments" (p. 1064). A similar conclusion was made in a recent review of CBAs in science and medical education (Kuo & Wu, 2013); among the 66 CBAs reviewed, only 19 incorporated dynamic or interactive media in item presentation. There is still a need for more research on CBAs that exploits the potential of computer technology to effectively tap into complex abilities in science.

Additionally, multiple forms of knowledge representations such as diagrams, animations, maps, graphs could play a crucial role in the design of CBAs (Mislevy et al., 2010) and have been frequently incorporated into CBAs for item presentation. These multiple external representations (MERs) have been widely used for teaching and learning (see reviews in Ainsworth, 1999; Wu & Puntambekar, 2012). To fully explore the potential of MERs, educational researchers have identified various design parameters and representational characteristics of MERs such as forms, sequences, and modalities (Ainsworth, 2006), and examined the effects of these parameters and characteristics on learning by integrating them into instructional materials (e.g., Mayer, Hegarty, Mayer, & Campbell, 2005; Wu, Lin, & Hsu, 2013). However, what remains relatively little understood is the effects of the use of MERs on students' performances in computer-based assessments, particularly in those evaluating complex science abilities.

Having considered the potential of CBAs in assessing complex abilities, therefore, we designed a multimedia-based assessment for secondary school students that focused on inquiry abilities in science education (Wu, Wu, & Hsu, 2014). To fill the gap of research on the effects of MERs in CBAs, through the design of test booklets and the participation of 1218 students (561 8th graders and 657 11th graders), this study aimed at investigating how modalities of representations (i.e., dynamic and static) and types of visual information (i.e., context and content) affected the item difficulty of the multimedia-based assessment. Three research questions were explored:

1. To what extent did the modality of representations for item presentation affect the item difficulty of the multimedia-based assessment?
2. To what extent did the modality effect vary by the student grade?
3. To what extent did the modality effect vary by the type of visual information?

In the following sections, we offer the theoretical background and empirical foundations of the study.

2. Literature review

2.1. A framework of inquiry abilities

Inquiry in science education can be defined as multifaceted activities similar to what scientists do to construct knowledge and propose explanations about the world (Duschl, Schweingruber, & Shouse, 2007; Krajcik & Czerniak, 2007; NRC, 2000). More specifically, these activities include "making observations; proposing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results" (NRC, 1996, p. 23). Engaging in the inquiry activities productively requires some cognitive abilities. NRC (2000) argued that these cognitive abilities were fundamental abilities necessary to do scientific inquiry; they should also go beyond science process skills (e.g., observation and inference) and "require students to mesh these processes with scientific knowledge as they use scientific reasoning and critical thinking to develop their understanding of science" (p. 18). Taking this integrated view of inquiry ability, this study also defined scientific inquiry abilities as the cognitive abilities to coordinate science knowledge and skills during engagement in relevant inquiry activities. Additionally, inquiry abilities have been emphasized in science education as essential competencies and as a part of scientific literacy (e.g., NRC, 1996, 2000; Wenning, 2007). This urges a need for valid assessments of students' inquiry abilities.

To design a multimedia-based assessment that effectively examined junior and senior high school students' inquiry abilities, we thus developed a framework of inquiry abilities in reference to the policy documents (MOE, 1999, 2008; NRC, 2000) and relevant research addressing scientific inquiry in classrooms (e.g., Krajcik et al., 1998; Krajcik & Czerniak, 2007; Nowak, Nehring, Tiemann, & Upmeier zu Belzen, 2013; Wu & Hsieh, 2006). Because it was not possible to evaluate all abilities mentioned in the literature in one assessment, after a careful consideration of the nature of inquiry, the scope of the assessment, the number of items, and the amount of testing time, we focused on four abilities and related sub-abilities, including questioning (e.g., asking and identifying questions), experimenting (e.g., identifying variables and planning experimental procedures), analyzing (e.g., identifying relevant data and transforming data), and explaining (e.g., making a claim and using evidence). The four abilities were selected to form our assessment framework because they were identified by NRC (2000) as fundamental abilities necessary to do scientific inquiry and were among the eight learning practices suggested by a new framework of K-12 science education (NRC, 2012). Both the new framework and our assessment shared focuses on asking questions, planning and carrying out investigations and analyzing and interpreting data, and constructing evidence-based explanations. In addition, our assessment framework subsumed the abilities to engage in arguments from evidence and to use mathematics and computational thinking in some sub-abilities. Furthermore, three levels of performance complexity were identified to capture major differences in student proficiencies (Table 1). This framework laid out a blueprint to promise a comprehensive coverage of the tasks across levels of performance complexity and across the inquiry abilities. Table 1 presents the inquiry abilities and examples of student performances at the three levels in our inquiry ability framework.