Success Factors for e-Assessment in Computer Science Education

Jirarat Sitthiworachart
School of Informatics
Walailak University
Thailand
sjirarat@wu.ac.th

Mike Joy
Department of Computer Science
University of Warwick
United Kingdom
M.S.Joy@warwick.ac.uk

Erkki Sutinen
Department of Computer Science
and Statistics
University of Joensuu, Finland
erkki.sutinen@cs.joensuu.fi

Abstract: Assessment plays a crucial role in improving learning outcomes, and e-assessment can be used to stimulate higher order thinking skills and provide personalized feedback automatically. In this paper, we review the uses of e-assessment within the domain of Computer Science Education, and identify the factors which are evidenced as contributing to the success of the e-assessment approach. An analysis of these data leads us to propose a novel framework for e-assessment which captures the essential success factors.

Introduction

Assessment is not only for the purpose of measuring students' knowledge and skills, but also for supporting educational goals (Bull & McKenna 2004 and Futurelab 2006). In traditional marking, students' assignments are marked by a course marker who may provide some feedback for the individual students' improvement, but this is a time-consuming task, especially for a large class. E-assessment is well accepted for automated marking with providing quick feedback repeatedly and consistently (Bull & McKenna 2004). Tselonis & Sargeant (2007) note that "marking is a dynamic process where the computer deals with repetitive tasks while the human makes the important judgments". E-assessment provides many potential benefits to students, including the following (Tselonis & Sargeant 2007 and Bull & McKenna 2004):

- monitoring student progress through frequent assessments;
- shortening the time gap between submission and receiving feedback;
- applying a variety of question types, for example using graphics and multimedia;
- promoting flexible learning and adaptive learning;
- monitoring question quality using statistical analysis;
- reducing the potential for cheating by randomizing questions; and
- sharing questions via question banks.

What has been done?

E-assessment research activities in Computer Science (CS) education in recent years are shown in Table 1. The main purpose of these e-assessment tools is supporting students in learning, for example identifying students' misconceptions and providing useful feedback for individual development (Kay *et al*. 2007 and Lilley *et al*. 2005). A variety of question types are used in e-assessment tools, such as multiple-choice questions (MCQs), text questions (including programming code), and diagram based questions (e.g. UML class diagram, conceptual database diagram). Major assessment techniques in CS education are summarized below and a comparison of main features of e-assessment systems are shown in Table 2.

- *Grouping answers* (both text and diagrams) can be used to improve speed and consistency when marking a large number of answers (Wood *et al*. 2006).
- *Diagram matching* against a model answer, or adoption of reverse and forward engineering techniques of consistency between diagram design and code structure, require more investigation, especially how to translate the results of matching into meaningful feedback. A new diagram editor is required to make the automatic assessment possible and also help students in diagram design (Hayes *et al*. 2007, Tselonis & Sargeant 2007 and Batmaz & Hinde 2006).

- *Self assessment and peer assessment* are effective techniques for supporting reflective learning and helping students to have a better understanding of assessment/course objectives (Kay *et al*. 2007 and Hamer *et al*. 2007).
- *Adaptive testing* is helpful for individual development (Lilley *et al*. 2005).

Authors	Subjects	Objectives	Question types	Assessment techniques/methods
Tselonis & Sargeant (2007)	Software engineering	To seek representations of answers and marking judgments which can be applied to a wide variety of situations	UML class diagram	• Matching the student's drawing answer against a model answer
Kay *et al*. (2007)	C programming	To identify and summarize students' common misconceptions in C programming	MCQ, Text	• Allowing the student to self-assess solutions; • Providing intelligent and informative learning progress feedback
Hamer *et al*. (2007)	Introductory programming	To investigate whether peer assessment can be an effective aid to learning	MCQ, Text	• Allocating submissions to reviewers, creating web-entries for reviews, accessing feedback, and calculating weighted-average grades
Hayes *et al*. (2007)	Object oriented development	To investigate the consistency between the student design and code	Diagram, Text	• Automatically assessing of the design-code interface (UML methodology and Java programming language)
Wood *et al*. (2006)	Artificial Intelligence	To develop deeper processing for fully automatic marking, using lightweight clustering	Text	• Calculating the similarity between short text answers and identifying clustering; • Providing per-cluster formative feedback
Batmaz & Hinde (2006)	Database design	To investigate semi-automatic assessment which helps the assessor by reducing the number of diagrams to be marked	Conceptual database diagram	• Grouping identical segments of the student's diagrams, so that the assessor can approve the correctness of a diagram fragment from the each of the different groups
Lilley *et al*. (2005)	Human computer interaction	To provide automated feedback for individual learners in a summative assessment context	Graph, MCQ (Computer-adaptive test: CAT)	• Using an adaptive algorithm based on the Three-Parameter Logistic (3-PL) Model from Item Response Theory; • Giving students feedback on: overall proficiency level, performance in each topic and recommended topics for revision

Table 1: e-Assessment systems in computer science education

In addition, problems with e-assessment in CS education are reported (Kay *et al*. 2007, Tselonis & Sargeant 2007, Bull & McKenna 2004 and Hamer *et al*. 2007), such as:

- e-assessment tools cannot be applied to all modules in CS;
- question types are limited – fully automated marking of essay and diagram based questions is a very difficult task, and semi-automatic marking might be applied;
- more intelligent feedback on individual learner's progress might be provided, and regular feedback to build student confidence is required;
- more research for individual development is required, for selecting the most suitable learning paths and appropriate choices of tasks/questions for individual students; and
- question quality cannot be measured accurately and cannot support evaluation of higher cognitive thinking skills (in Bloom's taxonomy).

e-Assessment systems main features	e1	e2	e3	e4	e5	e6	e7
	Gree (Tselonis & Sargeant, 2007)	Reflect (Kay *et al.*, 2007)	Aropa (Hammer *et al.*, 2007)	Design code interface (Hayes *et al.*, 2007)	ABC (Wood *et al.*, 2006)	Diagram marking (Batmaz & Hinde, 2006)	CAT (Lilley *et al.*, 2005)
Semi-automatic marking		/	/		/	/	
Fully automatic marking	/			/			/
Generate real time feedback	/			/			/
Provide feedback (not real time)		/	/		/	/	
Record learners' interactions with the system	/	/	/			/	
Relate to Bloom's taxonomy			/		/		/
Provide a model of learners progress/user profile (monitoring learners' learning progress)		/					
Provide learning guidance (adaptive advice for individual development)							/
Communication tool between author and reviewer			/				
Provide a diagram editor	/				/	/	
Describe learning objectives (domain knowledge)		/					
Automatic allocation of submissions to reviewers			/				
Weighted-average grade calculation			/				
Group similar answers					/	/	
Assess the consistency between the student code and design				/			
Adaptive tests							/
Question bank							/

Table 2: Main features of e-Assessment systems

Success Factors

From the review of e-assessment tools in CS education and the reports from several studies, we can classify success factors of e-assessment into 3 categories – human, tool designing, and technology (Dermo 2007, Ruedel *et al*. 2007 and QCA 2007).

Human

There are 3 groups of people which relate to the assessment process – teachers, students and examination boards. *Effective communication* among these people is perhaps the most important factor for the assessment process to be successful. The other factors of each group for success e-assessment are described in Table 3.

Teachers/ Staff	Availability	• Bertolo & Lambert (2007) report that students expected staff to be available 24/7, since the tests were continually accessible and they identify that establishing boundaries for staff availability is a prerequisite for a successful CAA exercise.
Students	IT skill	• Jones (2007) notes that if students do not possess basic IT competencies, effective e-learning cannot take place, and institutional support mechanisms may struggle to cope.
Examination boards	Agreement and support	• Ruedel *et al*. (2007) suggest that not only must examination boards consent to the use of e-assessment instead of traditional assessment in formative/summative, but also that they should provide technological support and guidance for its deployment.

Table 3: Human factors in e-assessment

Tool Designing

The design of an e-assessment tool is the most important stage. Tasks/questions should be suitable for each individual student and can be used to assess all skill levels (as defined in Bloom's taxonomy). Students must have a clear understanding of assessment objectives and the e-assessment process. Success factors of e-assessment tool design can be summarized in Table 4, with support evidences from 7 e-assessment systems (*e1-7*) in Table 2.

e-Assessment systems main features	e1	e2	e3	e4	e5	e6	e7
	Gree (Tselonis & Sargeant, 2007)	Reflect (Kay *et al.*, 2007)	Aropa (Hammer *et al.*, 2007)	Design code interface (Hayes *et al.*, 2007)	ABC (Wood *et al.*, 2006)	Diagram marking (Batmaz & Hinde, 2006)	CAT (Lilley *et al.*, 2005)
Semi-automatic marking		/	/		/	/	
Fully automatic marking	/			/			/
Generate real time feedback	/			/			/
Provide feedback (not real time)		/	/		/	/	
Record learners' interactions with the system	/	/	/			/	
Relate to Bloom's taxonomy			/		/		/
Provide a model of learners progress/user profile (monitoring learners' learning progress)		/					
Provide learning guidance (adaptive advice for individual development)							/
Communication tool between author and reviewer			/				
Provide a diagram editor	/				/	/	
Describe learning objectives (domain knowledge)		/					
Automatic allocation of submissions to reviewers			/				
Weighted-average grade calculation			/				
Group similar answers					/	/	
Assess the consistency between the student code and design				/			
Adaptive tests							/
Question bank							/

Tasks/ Questions	f1	Skill levels (as defined in Bloom's taxonomy): *e3, e5, e7*	• Chew & Jones (2007) report that although most assessment questions require learners to think only at a lower level, e-assessment can potentially be used to stimulate higher order thinking skills. • Farrell (2006) suggests that an author should actively think about Bloom's taxonomy when setting questions.
	f2	Adaptive test: *e7*	• Barker & Lilley (2006) propose using adaptive testing as a basis for automated feedback to target individual learners' strengths and weaknesses in specific topics. • Winkley & Osborne (2006) note that a system's adaptivity towards learner activities speeds up formative assessment and is a strong motivating factor.
Objectives	f3	Clear objectives: *e2*	• Kay *et al.* (2007) suggest that students will be better able to evaluate example solutions and write their own if they have previously understood the learning objectives to which the learning task relates.
Process	f4	Consistency and reliability	• QCA (2007) research evidences "demonstrable consistency and reliability" as being a critical success factors.

	f5	Interactive and multimedia: *e1, e5, e6*	• Ekins (2007) reports that interactive quizzes are not only enjoyable and stimulating for students, but also help them to check whether they have understood their learning material.
	f6	Feedback: *e1-7*	• Hattie (1987) identifies good quality feedback as the stated that "most single powerful influence on student achievement in higher education."
	f7	Monitor: *e2*	• Kay *et al*. (2007) proposes a model of learner's progress in reflection learning (self assessment) by monitoring learners' learning progress and summarizing learners' misconceptions.
	f8	Report: *e1-3, e6*	• Ekins (2007) notes that a log of students' responses, scores, time spend on each question, etc., are data which can be used to analyze student errors.
	f9	Help	• Dermo (2007) suggest that support plans and help facilities are required components of an e-assessment system.
Others	f10	Bias	• Hamer *et al*. (2007) and Davis (2003) suggest that anonymous interaction should be applied to avoid individual markers' bias.
	f11	Security	• Whitelock & Brasher (2006) identify confidence in the security of delivering marking e-assessment assignments as of importance, especially for large scale assessments.

Table 4: Tool designing factors in e-assessment

Technology

As technology changes rapidly, e-assessment should be updated (both hardware and software). Many problems might arise while keeping up with technology, such as cost and quality, and hardware and software need to be carefully monitored to avoid system failure.

A Framework for e-Assessment

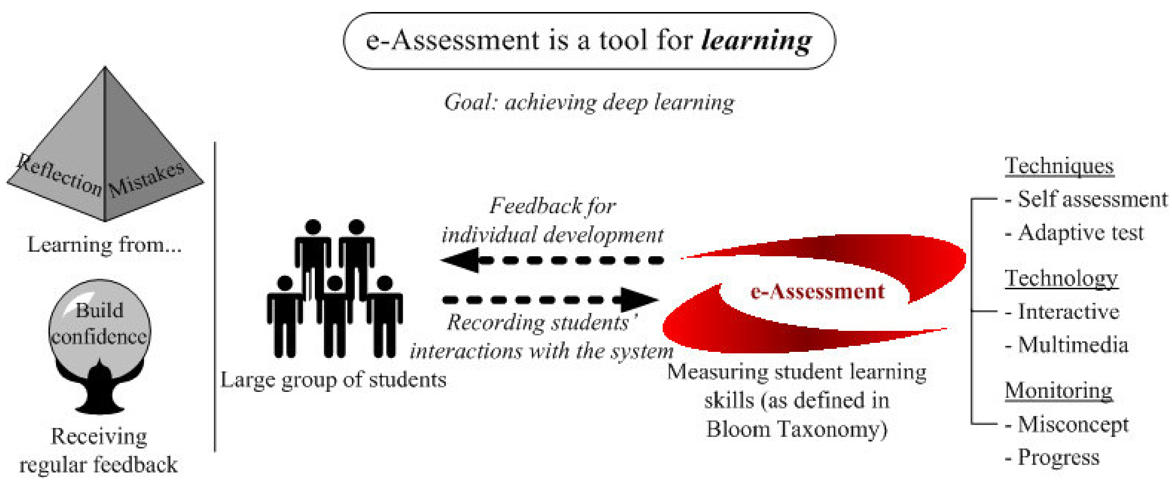

Figure 1: A framework for e-assessment (an intelligent adaptive deep learning tool)

An analysis of these data leads us to propose a novel framework (an intelligent adaptive deep learning tool) for e-assessment which captures the essential success factors (*f1, f2, f5-f8* in Table 4). In order to improve learning outcomes in higher education, e-assessment can be used to stimulate the process of evaluation, teaching and learning (Futurelab 2006). Hamer *et al*. (2007) warn us that "it has to be done carefully; it needs to be approached not as an evaluative tool but pitched as away of learning". Student thinking skills, as defined in Bloom's taxonomy (Bloom *et al*. 1956), should be considered when designing an assessment tool (Chew & Jones 2007 and Farrell 2006). Most assessment tools evaluate only the lower thinking level (i.e. knowledge, comprehension, and application), because it is not easy to set exam questions that measure deep learning – analysis, synthesis, and evaluation (Guest & Brown 2007 and Chew & Jones 2007).

Figure 1 displays a framework for e-assessment that encourages higher cognitive thinking skills which support an individual student's development, which includes the following essential factors.

- *Techniques (f1, f2)*: self assessment and computer adaptive testing (ranking questions based on Bloom's taxonomy) represent an advance in educational assessment which motivates individual students appropriately by providing personalized feedback (Wheadon & He 2006 and Barker & Lilley 2006).
- *Technology (f5)*: interactive and rich multimedia adds value to an e-assessment system and can present complex data in ways that are easy to understand.
- *Monitoring (f7)*: e-assessment is useful for diagnosing student misconceptions and monitoring students' progress.
- *Feedback for individual development (f6)*: successful feedback should focus on learning (rather than on marks) in order to improve students' learning, and should be clear and understandable to each student (Bedford & Price 2007).
- *Recording students' interactions with the system (f8)*: a log of students' responses is useful for analyzing student errors (Ekins 2007).
- *Learning from reflection and mistakes (f7)*: students potentially have much to learn from critically reflecting on the mistakes they have made in self assessment (Kay *et al*. 2007).
- *Receiving regular feedback (f6):* regular feedback builds confidence and this improves students' learning (Bull & McKenna 2004).

Conclusion

We have reviewed the uses of e-assessment within the domain of Computer Science Education, and have identified the factors which are evidenced as contributing to the success of the e-assessment approach. An analysis of these data leads us to propose a novel framework for e-assessment which captures the essential success factors. We are currently designing and developing an e-assessment tool for CS education to foster deep learning and focus on providing an intelligent personalized feedback automatically, which might be described as 'an intelligent adaptive deep learning tool'.

References

Barker, T. & Lilley, M. (2006). Measuring Staff Attitude to an Automated Feedback System based on a Computer Adaptive Test. *International Computer Assisted Assessment Conference 2006*, Loughborough University, UK.

Batmaz, F. & Hinde, C. (2006). A Diagram Drawing Tool for Semi–Automatic Assessment of Conceptual Database Diagrams. *International Computer Assisted Assessment Conference 2006*, Loughborough University, UK.

Bedford, S. & Price, G. (2007). A Study into the Use of Computer Aided Assessment to Enhance Formative Assessment during the Early Stages of Undergraduate Chemistry Courses. *11th CAA: International Computer Assisted Assessment Conference 2007*, Loughborough University, UK.

Bertolo, E. & Lambert, G. (2007). Implementing CAA in Chemistry: A Case Study. *11th CAA: International Computer Assisted Assessment Conference 2007*, Loughborough University, UK.

Bloom, B.S. (1956). Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain. New York, Longmans.

Bull, J. & McKenna, C. (2004). *Blueprint for Computer-Assisted Assessment*. RoutledgeFalmer, London.

Chew, E. & Jones, N. (2007). The Marriage of Freire and Bloom: An Assessment Prototype for Pedagogy of the Oppressed and Higher Order Thinking. *International Computer Assisted Assessment Conference*, UK.

Dermo, J. (2007). Benefits and Obstacles: Factors Affecting the Uptake of CAA in Undergraduate Courses. *11th CAA: International Computer Assisted Assessment Conference 2007*, Loughborough University, UK.

Ekins, J. (2007). The Use of Interactive on-line Formative Quizzes in Mathematics. *11th CAA: International Computer Assisted Assessment Conference 2007*, Loughborough University, UK.

Farrell, G. (2006). A Comparison of an Innovative Web-based Assessment Tool Utilizing Confidence Measurement to the Traditional Multiple Choice, Short Answer and Problem Solving Questions. *10th CAA: International Computer Assisted Assessment Conference 2006*, Loughborough University, UK.

Futurelab (2006). *Report 10: Literature Review of E-assessment*. [online] Available from: http://www.futurelab.org.uk/research/lit_reviews.htm (accessed March 25th 2008).

Guest, E. & Brown, S. (2007). A New Method for Parsing Student Text to support Computer-Assisted Assessment of Free Text Answers. *International Computer Assisted Assessment Conference 2007*, UK.

Hamer, J., Kell, C. & Spence, F. (2007). Peer Assessment Using Aropa. *Ninth Australasian Computing Education Conference (ACE2007)*, Ballarat, Australia.

Hattie, J. (1987). Identifying the salient facets of a model of student learning: a synthesis of meta-analyses, *International Journal of Educational Research*, *11*, 187-212.

Hayes, A., Thomas, P., Smith, N. & Waugh, K. (2007). An Investigation into the Automated Assessment of the Design-Code Interface. *Conference on the Innovation and Technology in Computer Science Education*, UK.

Jones, S. (2007). Diagnosing and Developing the IT Skills of New Entrants to Higher Education. *11th CAA: International Computer Assisted Assessment Conference 2007*, Loughborough University, UK.

Kay, J., Li, L. & Fekete, A. (2007). Learner Reflection in Student Self-assessment. *Ninth Australasian Computing Education Conference (ACE2007)*, Ballarat, Australia.

Lilley, M., Barker, T. & Britton, C. (2005). The Generation of Automated Learner Feedback Based on Individual Proficiency Levels. *18th international conference on Innovations in Applied Artificial Intelligence*, Bari, Italy.

QCA: Qualifications and Curriculum Authority. (2007). *e-Assessment: guide to effective practice*. [online] Available from: http://www.efutures.org/docs/guide.pdf (accessed March 25th 2008).

Ruedel, C., Whitelock, D. & Mackenzie, D. (2007). Key Factors for Effective Organisation of E-Assessment. *11th CAA: International Computer Assisted Assessment Conference 2007*, Loughborough University, UK.

Tselonis, C. & Sargeant, J. (2007). Domain-Specific Formative Feedback through Domain-Independent Diagram Matching. *International Computer Assisted Assessment Conference 2007*, Loughborough University, UK.

Wheadon, C. & He, Q. (2006). An Investigation of the Response Time for Maths Items in a Computer Adaptive Test. *10th CAA: International Computer Assisted Assessment Conference 2006*, Loughborough University, UK.

Whitelock, D. & Brasher, A. (2006). Developing a Roadmap for e-Assessment: Which Way Now?. *10th CAA: International Computer Assisted Assessment Conference 2006*, Loughborough University, UK.

Winkley, J. & Osborne, C. (2006). Developments in On-Screen Assessment Design for Examinations. *CAA: International Computer Assisted Assessment Conference 2006*, Loughborough University, UK.

Wood, M., Jones, C., Sargeant, J. & Reed, P. (2006). Light-weight Clustering Techniques for Short Text Answers in Human Computer Collaborative (HCC) CAA. *International Computer Assisted Assessment Conference 2006*, UK.