ONLINE ASSESSMENT OF PRACTICAL EXPERIMENTS

Ruth Thomas and Colin Milligan

Online Assessment of Practical Experiments

Dr Ruth C Thomas & Dr Colin Milligan
SCROLLA, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS.
Email: rutht@ma.hw.ac.uk, colinm@ma.hw.ac.uk

Abstract

Online assessment has yet to make much progress into the testing of higher order skills. One way to accomplish this is to use a form of authentic assessment in which assessment occurs in the environment in which the student has learnt or is learning. This paper explores the potential for such an approach in a simulated practical experiment.

The type of learning which can occur within a simulation and which we ultimately wish to assess is wide-ranging. Simulations can be used to provide students with an environment in which they can explore or undertake investigative, problem or task-based learning.

A form of intrinsic feedback occurs naturally in simulations in that the learner sees the results of any action taken. However, to provide formative assessment to assist learning, further feedback is required. Our premise is that, a more flexible and effective system can be provided by integrating the system capable of building such simulations (JeLSIM toolkit) with an assessment engine, initially CUE.

We see the role of the teacher in a classroom as providing a model of authentic assessment that the computer-based system must be able to duplicate. We have constructed a simple exemplar system from existing simulation and assessment resources, to inform the design of the more fully functional system.

The more fully functional system would have the potential to unobtrusively test learning outcomes, such as the students' ability to plot graphs correctly, ask students to make complex predictions and test their accuracy, assess the student's ability in designing, planning and running an experiment, or to assess the student making and testing hypotheses.

Keywords: Authentic assessment, Simulations, Laboratory Experiments, Constructivism, JeLSIM, CUE.

Introduction

The constructivist view of assessment is that it should be authentic and interwoven with the education process (e.g. Jonassen 1991, Jonassen 1994). That is, it should be ongoing and assess the overall learning process as well as the final product. This idea of "authentic" assessment that takes place in the environment in which the student has learnt or is learning, is the focus of this paper; our interest being in the area of Computer Aided Authentic Assessment (CAAA). Our ultimate aim is to design and build a system capable of undertaking authentic assessment of computer-based tasks. In this paper we explore the potential of this approach in a simulated practical experiment.

Our initial work was to review the teachers' role in a traditional laboratory setting. An understanding of how the teacher uses authentic assessment is crucial in defining the requirements of a computer-based system, which can be used where the teacher is either unable to deal with all students individually, or is not present at all.

Our premise is that a starting point for a computer-based system is to look at ways of utilising assessment engine technology with an online task such as a simulated practical experiment. Our first step has been to produce a simple prototype demonstrator using a simulated practical experiment and the CUE assessment engine. This prototype has highlighted a number of areas where fuller integration is required to achieve a more flexible system.

Finally, we explore the functionality and potential of a more fully integrated system in assessing online practical experiments

Simulations as tasks for authentic assessment

Simulations are an ideal computer based resource to use as a task for authentic assessment because they have a unique role in supporting learning Pilkington and Parker-Jones (1996) and can be used to provide realistic activities, of some duration and depth.

Simulations allow learners to directly manipulate a system and to observe the effect of the change, providing the learner with a form of intrinsic feedback. In practice, this feedback is not sufficient to allow most students to learn from a simulation and the necessity for additional guidance, feedback and scaffolding as a student uses simulations has been recognised for some time (Thomas and Neilson 1995). Some support and guidance can be provided by linking the simulation with other multi-media resources, which set the scene or provide answers to frequently asked questions, but in a typical simulation, feedback related to the students' immediate need or current task, or assessing their performance, is limited or absent. For more complicated tasks, explanation and advice will be provided by a teacher or demonstrator.

There are two important factors to consider when choosing a task for authentic assessment:

1. The task must be realistic. It should have meaning to the student outside of the computer-based environment.
2. The learner should engage with subject matter not the software. It is vitally important that the system does not end up assessing the student's ability to press buttons in the right order.

In this work we have chosen to look at the assessment of simulated practical experiments as they provide a computer-based activity of reasonably long duration that involves a range of different skills.

The role of teacher in laboratory and computer-based practicals

A starting point for computer-based authentic assessment is to attempt to duplicate the role of a teacher in the classroom. An important component of formative assessment in teaching is the feedback (Laurillard, 1993):

- from the teacher to students to help learners become aware of any gaps in their knowledge, understanding, or skill and to guide them to overcome difficulties; and
- from students to the teacher as they undertake a task, allowing the teacher to assess the learner's understanding and teachers to modify their teaching strategies accordingly.

These feedback activities are ongoing throughout the assessment process and a computer-based equivalent must be able to duplicate them.

We have examined the role of the teacher in providing feedback in laboratory practicals (both traditional and lab-based) to inform the requirements for a CAAA approach. The teacher provides two forms of feedback:

1. Solicited where the learner initiates the request for assistance to ask for:
    - instructions or information,
    - help when they know they have made a mistake or are stuck,
    - reassurance on the validity of an action,
    - an explanation of an observation,
    - confirmation of their explanation of a phenomenon.
2. Unsolicited, where the teacher notices that a student or the entire class is in difficulty and they may:
    - modify the task,
    - provide more explanation or resources,
    - provide contingent tutoring (Wood and Wood, 1999),
    - mark the activity.

The ideal computer aided authentic assessment system ought to be able to duplicate this behaviour.

Combining simulation and assessment engine

Simulations can be used effectively when external expertise of some form can be provided. When the teacher is not present, it is necessary to provide feedback in some other way. It is possible to use a programmer to hardcode feedback and assessment into a simulation, however, feedback is usually limited to the provision of generic errors and warning messages for two reasons:

1. a single simulation can be used in many different teaching contexts and levels and hard-coded feedback would not have universal relevance.
2. feedback and assessment should be under the control of the experts, i.e. the teachers, and its provision should not require programming skills.

Rather than hard coding such feedback a much more flexible and effective system can perhaps be provided by integrating simulations with an assessment engine.

The JeLSIM tools (Thomas et al 2000[1] and URL: 1) are ideal for the simulation component of the system, because they allow a non-programmer to construct a specific visualisation of a simulation model. A single model can give rise to many visualisations, which can be saved and delivered as Java applets. Thus it is easy to modify existing materials to suit different learning contexts and it would be possible to build new questions from existing material with very little additional effort.

Ultimately the combined simulation-assessment system will work with any assessment engine. We've chosen to use CUE (Paterson, 2002) as we work closely with the developers and it is one of the more sophisticated engines available.

The first stage in our investigation has been to construct an exemplar of the idea from existing components. This exemplar has been used as to inform the requirements analysis of our design of a more sophisticated integrated system.

A simple prototype system

We have built a simple system for assessing a simulated practical from two pre-existing components:

1. The CUE assessment engine. (CUE has been enhanced to allow the inclusion of a Java applet within a question.)
2. A simulated practical based on one of a series simulated chemistry practical experiments, produced by JeLSIM Partners for the Scottish Qualifications Agency (SQA) and the University of Cambridge Local Examinations Syndicate (UCLES) for the project *"Simulating online experiments"* (URL: 2). The resources deal with the subject of chemical kinetics (the study of factors affecting the rate of chemical reactions).

[1] This reference refers to the functionality of the MultiVerse toolset. The JELSIM tools are based on the MultiVerse toolset and have similar functionality.

The idea in building this prototype was not to produce something suitable for use in a classroom, but to explore the range and limitations of a system built on current technology.

Overview of the simulation task

Part of the *"Simulating online experiments"* project involved a review of the concerns of teachers over the use of online practicals, as a result, the simulated experiments do not attempt to mimic the mechanics of an experiment and the dexterity required to perform it, but rather attempt to get the student to duplicate the thought processes that are required to complete the experiment. The aim is not to replace, all practical experiments with a computer-based equivalent, but to augment existing resources in these areas by using them as:

- a variant of the real experiment to avoid direct replacement,
- preparation for the student in advance of the laboratory exercise,
- revision resources,
- a way of providing activities in experimental design,
- support material for subject teaching in chemical kinetics.

The resources are designed to be realistic rather than paint an idealised picture of the experiment:

- students collect data which is subject to experimental error and they can make mistakes and get erroneous results,
- phenomena that happen in the computerised experiment take the same length of time as in the real world.

The applet we have used in the exemplar is a simulation of an experiment to determine the effect of concentration on reaction rate. It consists of 5 tabbed screens:

1. **Introduction** – Description of the experiment
2. **Prepare mixtures** - where the learner mixes chemicals together in different amounts to produce solutions of various concentrations,
3. **Run experiment** - where the learner starts the reaction and measures the time until the solutions change colour,
4. **Prepare graph** - where the learner decides what to plot and the units and range for each graph axis,
5. **Plot graph** - where the learner examines the results and draws conclusions about the experiment.

Implementation

When the exemplar is running, CUE and the JeLSIM applet sit side by side on a vertically split screen.

The CUE window consists of a single, twelve item, multipart, question on a single page. Learners must scroll down to see all the questions. Any feedback appears at the top of the CUE application in a separate window.

The main question at the top of the CUE screen contains a small applet that the learner can use to launch the simulated task applet in a separate web browser.

The applet window consists of a modified version of one of the experiments from the "Simulated Practical Experiments" project. Students work through the experiment as they would normally but the simulation interface has been amended to prompt the user to go to CUE to answer a series of questions when they complete each stage of the experiment.

Examples of the types of questions asked are given below:

- On the prepare mixtures screen:
- To check that the user has chosen sensible values for the volumes in each mixture.
- On the run experiment screen:
- To check that learners know how to calculate the concentration of a chemical,
- To check that their readings of times for colour change are sensible.
- On the prepare graphs screen
- To check that they can calculate reaction rate,
- To check the units they chose for each axis,
- To check the maximum and minimum plot range they chose on an axis.
- On the plot graphs screen
- To check that their data gave the expected results and they could appreciate the trend of the data,
- To check that they could draw a conclusion from the experiment.

In each case feedback is given once the learner submits an answer.

The good features of the implementation

It proved possible to develop a series of questions (multiple choice and judged mathematical expressions) linked to a single task of some 30 minutes duration and to display them in a usable format. Questions were authored that were relevant to each stage of the simulated practical and which could potentially check that whether or not the student was having problems with the exercise. A number of the questions are of the type teachers ask if they are assessing the student's understanding or checking that the student hasn't made a mistake early in the experiment which will invalidate the rest of the experiment, for example, "How did you calculate that value?".

The feedback feature of the assessment engine has been used to either give them details about an error, or ask them a question to prompt them to realise they have made a mistake and remedy it themselves, much as a teacher would.

Limitations of the implementation

Probably the biggest limitation of this exemplar in creating an authentic assessment system is the intrusive nature of the assessment. There was no communication between the simulation and the assessment engine, so that in a number of cases, the student had to perform a task in the simulation and repeat that task in CUE in answer to a question. For instance, when they choose the units for the axes on the graph they wish to plot, they have to make exactly the same choice in CUE again. This is almost the equivalent of a teacher badgering a student with questions all the time as they undertake a task and is clearly not ideal.

This problem would not arise in a fully integrated system where simulation variables could be passed directly to the assessment engine. Under ideal circumstances, the assessment engine would receive information from the simulation, but only intervene if the student requires assistance, in the same way a teacher would in a laboratory exercise.

In the exemplar, a large number of questions are multiple choice where ideally they might have been better posed as questions requiring a numeric answer. The reason for this is that, in simulations, the learner is given freedom of choice and there are many valid paths which can be taken through a simulation so it is often not possible to know in advance what value a variable will have at a certain point in a task. For example in the practical experiment, the student is monitoring the time it takes for a colour change to appear in a solution. The reading they take will depend not only on the concentration of the solution, but also on how quickly they started the stopwatch when they mixed the chemicals, how soon they noticed the colour change and how quickly they stopped the clock. It is not therefore possible to check such an answer with an assessment engine as a numeric without information about other variable values in the simulation. In this simple exemplar therefore we used a multiple choice that asked whether their reading fell within certain ranges of time and gave feedback according to their choice. In a system where the simulation could communicate with the assessment engine, it could pass out information about the time that the concentration of a chemical reached a certain level and triggered the colour change and give the learner feedback as the accuracy of their measurement.

In order to produce a viable example of a question with a numeric answer, the flexibility of the simulation was decreased to ensure that the learners used a particular starting value so a correct answer would be known. This clearly isn't making best use of the simulation and would not be necessary in a more fully integrated system.

Another problem is screen “real estate”. This is going to be a problem whenever two large applications run together. Considering this from a usability standpoint, it is important that the user can see the task all the time (I.e. it does not scroll off the screen or disappear under other windows). It may be that ultimately the use of assessment engine technology for authentic assessment does not require the user to see the assessment in the form of a separate application running a test, but just to appear when required as the ideal teacher might in similar circumstances.

To overcome these limitations, a more fully integrated system must at the very least have some way of communicating information about simulation variables to the assessment engine.

An integrated system

Integration of assessment engine and simulation is a major project. Ideally it would involve providing the ability to pass information about simulation variables and their values between the simulation and assessment engines. This would allow the assessment engine:

- to define the value of any of the simulation variables when the simulation is started; and
- to receive information about any simulation variables as the student works on the assessed task.

The biggest difference between the exemplar system and a fully integrated system would be the ability to assess any numeric aspect of the task as the learner was doing it, but without the learner breaking off to enter values into an assessment engine. The type of assessment carried out, may well have the same intentions as in the exemplar, but they will be much less cumbersome and will only result in feedback to the learner if there is a problem.

The information contained in the simulation, both simulation variables and student input values, would be available at the assessment engine. Thus it would not matter what choices the student made in the running the simulation, the assessment engine would be able to assess how accurately students have calculated the values from **their own** experimental data.

Combination of such assessments throughout a practical exercise could begin to provide authentic assessment of the simulated experimental process.
The sophistication of feedback and assessment that could be provided in such a system obviously depends on the assessment engine, those with features such as randomisation, questions with multiple parts and steps could provide greater flexibility.

Other forms of assessment

It may be possible to collect data from the simulation:

- at regular pre-defined intervals; and
- at any time a variable is changed, (providing a history of student activity)

Given the availability of such data, it is possible to envisage tools that might intervene with hints, suggestions or contingent questions depending on student actions, much as the teacher would in such a situation. Such data could also provide information for debriefing students after exercises or to give the teacher insight of the strategies adopted by the student in tackling a problem. In such a system, it may well be possible to distinguish a problem-solving attempt involving reflection from one involving random guesswork. Ultimately, we would hope to be able diagnose student misconceptions and to provide learners with appropriate material to remediate the misconception.

Questions arising

This paper has focussed on formative assessment, but is there actually a difference between formative and summative assessment in an authentic assessment process? Or can summative assessment take place in the same system with reduced feedback to the learner? One of the educational benefits of using a computer-based simulation is the possibility of exploration, hypothesis and learning through mistakes. In summative assessment, should the student be rewarded or penalised for exploration? Should students be given more than one attempt at a task to encourage learning?

Another issue is whether, in authentic assessment, the student should be aware that assessment is occurring. Should assessment engine technology be used, but be hidden from the learner. In such a case answers could perhaps be evaluated, and the user given feedback and even a breakdown of performance, but without it taking the form of a traditional test.

Summary

It is clear that a combination of simulation and assessment engine could provide considerable advances on current practice in computer-aided assessment. It could allow the production of more complex question types that test deeper understanding. Potentially it would allow computer aided authentic assessment of practical experiments, which could reduce costs and pressure on resources.

However, simply having the technology does not guarantee the production of educationally sound questions. Teachers, subject experts and educationalists must be in control of the process. A key feature of the proposed system is that once the initial simulation model is written, a programmer is not required to produce these sophisticated questions.

Acknowledgements

The software used in the production of the exemplar was based on work from the project "Simulating online practical experiments" carried out by JeLSIM Partners and funded by the Scottish Qualifications Agency and the University of Cambridge Local Examinations Syndicate.

References

Jonassen, D. (1991) *Evaluating Constructivist Learning.* Educational Technology **31** (9) 28-33.

Jonassen, D. (1994) *Thinking Technology: Towards a Constructivist Design Model.* Educational Technology **34** (4) 34-37

Laurillard, D. (1993) *Rethinking University Teaching: A Framework for the Effective Use of Educational Technology*. London:Routledge.

Paterson (2002) *The CUE assessment system*.
<http://ltsn.mathstore.ac.uk/articles/maths-caa-series/apr2002/index.shtml>

Pilkington, R. and Parker-Jones, C. (1996) *Interacting with Computer Based Simulation: The Role of Dialogue.* Computers and Education **27** (1) 1-14

Thomas, R. and Neilson, I.(1995) *Harnessing Simulations in the Service of Education: The Interact Simulation Environment.* Computers and Education **25** (1/2) 25-29

Thomas, R., Close, A. and McAndrew, P (2000) *Tailoring Simulations for Teaching and Learning: The Potential of the MultiVerse Environment.* International Journal of Electrical Engineering Education ***37*** 1-2

Wood and Wood (1999) *Help Seeking, Learning and Contingent Tutoring.* Computers and Education **33** 153-169

URL: 1 *The JeLSIM Tools*. <http://www.jelsim.org>

URL:2 *Simulating Online Practical Experiments*.
<http://www.jelsim.org/pub/SQA-UCLESWebresource>