SIMASSESSMENT: ENHANCING ACADEMICS' UNDERSTANDING OF ASSESSMENT THROUGH COMPUTER SIMULATION

Shirley Alexander, Peter Kandlbinder, Elisabeth Howson,
Lawrence Lukito, Andrew Francois & Simon Housego
Institute for Interactive Media and Learning
University of Technology, Sydney, AUSTRALIA
Shirley.Alexander@uts.edu.au
Peter.Kandlbinder@uts.edu.au
Elisabeth.Howson@uts.edu.au
Lawrence.Lukito-1@uts.edu.au
Andrew.Francois@uts.edu.au
Simon.Housego@uts.edu.au

Abstract

Assessment of students is commonly seen as having two purposes: providing students with feedback on their progress (formative), and making judgements about an individual student's fulfilment of subject objectives (summative). Not so commonly understood by academics however, is the degree to which students' perception of assessment drives their learning. Many years of research have highlighted the fact that students respond to their perception of assessment in terms of what they learn and in the way in which they approach their learning. These results have been discussed in a plethora of conference papers, journal articles and books devoted to good practice in assessment. This work however, goes largely unnoticed by the majority of academics, many of whom continue to design assessment tasks which mirror their own experiences as students of large summative examinations.

This paper describes a project which seeks to draw academics' attention to good practice in assessment and to enhance their understanding of the impact of a range of issues related to assessment design on student behaviour and approaches to learning. The authors have developed an online assessment simulator, so that academics can try out different assessment approaches, and see first hand the ways in which students might respond. The simulator also demonstrates the links between assessment design, and student learning and behaviour.

Keywords
Assessment, simulation, quality learning, higher education, formative evaluation

The problem of assessment

Governments of many countries have highlighted assessment as a critically important issue in improving the quality of higher education. Within Australia, two of the recommendations of the recent Senate inquiry in to higher education were related to improving assessment. Recommendation 14 proposes that the Australian Universities Quality Agency (AUQA) "address the issue of course assessment to ensure the integrity of qualifications granted by Australian universities". Recommendation 15 advises universities to consider the more widespread use of external examiners and the greater use of moderation across a number of universities. As well as governments, students are also active in highlighting assessment as a priority area for major improvements.

Entwistle (1995) cites a study by Pennington (1994) which reported on the analysis of 600 feedback questionnaires at the University of Teeside. Half the students surveyed reported the need for more effective teaching delivery, twenty three percent wanted greater coherence in assessment methods, and sixteen percent, better course management and design.

Locally, in an analysis of 3,200 written comments by students of the University of Technology, Sydney (UTS) in the Course Experience Questionnaire, Scott (2002) reported that approximately 20% of comments related to assessment. Specifically, the common areas cited as needing improvement were:

- provision of clear guidelines, standards and expectations;
- consistency in marking (especially where group work is involved);
- prompt focused feedback on work; and
- the need to reduce the number of assessment tasks students are asked to complete.

In another study at UTS Bamford (2001) interviewed 27 students and 60 academic staff, and found that "…students complain of unclear tasks, increasing workload, and poor feedback. Students see assessment as a method of control, and lecturers tend to propagate this view through the use of assessment to regulate student behaviour".

At the same time Bamford reported that "Lecturers express concern about the perceived lowering of standards, the increased incidence of cheating and plagiarism and the pressures of growing class sizes and assessment overload".

The reports described above are just three of a large number of similar studies conducted in many countries, all of which highlight the high degree of similarity in problems experienced by students in different countries. Resolution of these problems should not be difficult as much is known about what constitutes good practice in assessment. However, what still eludes us is a more effective process for disseminating that knowledge in such a way that assessment practices are improved.

The problem of enhancing assessment

Previous attempts to improve the quality of learning by enhancing the quality of assessment have included provision of workshops, seminars, formal award courses, university-wide forums on assessment, distribution of reports on good practice, and distribution of printed and web-based materials on good practice in assessment.

However, the problem of improving practice is not a simple one as a number of factors create disincentives for academics to experiment with assessment. Students are highly sensitive to changes in assessment procedures and, as has been noted, assessment remains a major area of complaint by students when asked about the quality of university teaching. Given the need for transparency in assessment, most universities have regulatory requirements that forbid changes to the published assessment regimes once the semester begins. Further, certain practices such as end of semester examinations, have so dominated the university landscape that, as students, academics have only rarely experienced alternative assessment processes themselves. With so much at stake in assessment it appears reasonable to them to retain traditional assessment approaches regardless of the consequences to student motivation and learning.

As a consequence the classroom is an unfavourable environment for academic staff to experiment with alternative approaches to assessment. Recognising that different people like different approaches to engaging with their professional development, the Institute for Interactive Media and Learning (IML) has undertaken the development of an assessment simulator to encourage academic staff to review their assessment practices. The assessment simulator is an opportunity for them to implement changes in an assessment pattern and to get an approximation of the ways in which quite simple changes in the assessment of a subject can affect the way students engage with the subject content. The simulator provides academics with an indication of the likely impact of their choices in regards to students' approaches to learning, their motivation and emotional state.

SimAssessment - An assessment simulator for academics

Simulators of academic practices are not common in academic development. What examples there are, such as Entwistle et al (1977), generally target the students rather than lecturers. At the time this project was conceived, a number of competing strategies were identified including:

- an "Assessment Wizard" which would lead academics through the process of creating assessment items;
- a "quick Quiz" (3 minutes) to determine a user's assessment profile, leading to a 'remediation' pathway through an "Improving Assessment" website according to the profile developed;
- a website that provides opportunities for users to critique the assessment items developed by others;
- a "snakes and ladders" game where dice can't be thrown until a question on assessment has been answered correctly.

To ensure the prospect that academics will adopt the final project a focus group of six academics from a range of disciplines met to workshop these proposals and determine priorities for development. This group selected the SimAssessment project as the first priority for development. Participants in the focus group felt that a simulation that offers realistic scenarios would permit a wider range of approaches to assessment. In this way the simulator could provide feedback and a successful formula for assessment that could be used in the lecturer's class. The decision was made to proceed to develop a prototype of SimAssessment as the first priority, and the Assessment Wizard as the second priority.

Clearly, the primary goal of the assessment simulator was for academics to learn about the influence of different factors on students' approaches to learning.
Due to the innovative nature of the product it was initially decided to produce a prototype involving a subset of the full functionality so that extensive testing with potential users could be carried out prior to full development. The design phase involved the detailing of the different screens of the product that provided a "proof-of-concept". At that time, the design team was engaged in debate centred on the need to produce a full working prototype versus a series of flat screens that a user could step through. The final decision was made in favour of the former, in order to maximize the quality of information that could be gained from usability testing.

Concepts for the product were developed during the 'look and feel' design phase. In order to give the prototype a design that was as close to the real product as possible, student and mentor characters were also developed. These characters were eventually discarded for the usability test and a simplified approach to the characters was devised in order to focus the users' attention on the functionality. While the design was being developed, a specification for the backend approach was also being tested. A decision to develop SimAssessment for the web had already been made in order to provide maximum access for the audience. There was much debate over an object-oriented versus a rule-based approach to the programming, and a combination of the two was eventually decided upon.

During the production phase, the design was "cut up" for use in the Flash front end. The backend database was developed with input from the subject matter expert and an HTML front end was used initially by the programmer to replicate the Flash component. The back and front ends were not merged until the functionality to be tested in the prototype was finalized. A usability-testing plan was also devised to ensure that all aspects of the product were tested and which provided useful information for inclusion in the production of the final product.

Features of SimAssessment

The assessment simulator is comprised of four main components — registration, classroom interaction, debriefing and report. Once an academic logs in and nominates a subject area (figure 1), he/she is asked to choose a proposed assessment pattern for one semester of study. Alternatively, the user may run the simulation using the default values for the subject chosen. To guide their selections they have access to relevant documentation, such as the subject outline, last year's student feedback, and university policies relating to assessment. Additionally, they have the support of three software agents: the course

coordinator; a colleague; and an academic developer to choose an appropriate assessment pattern for the nominated subject.

Figure 1. simAssessment log-on screen

As each assessment item is added to the subject outline, the academic has to determine the goals of the assessment task, the type of assessment, how it relates to other assessment items in the semester, its weighting, who will do the marking and the type of feedback students will receive. The combination of these choices determines the responses of the other three components of the simulation.

The class action

The first version of the simulation takes place in a classroom of fifteen students, each with slightly different personalities and approaches to learning. Prosser and Trigwell (1999: 73) reported four student perceptions of their learning environment with four corresponding approaches to learning and this research provided the basis for the student characters. In a real class these characteristics would have been determined by the students' previous experiences with university teaching. In the simulation they are randomly assigned to provide a range of student responses as each of the student characters reacts to each assessment item as the semester progresses. In wanting academics to adopt an experimental attitude to their assessment choices, it is important that the simulation has a high degree of dramatisation. As a learning tool, the student responses were cues to the appropriateness of different assessment choices rather than the often difficult to read responses of real students. Being a cue to principles of assessment it was important the simulator allow the academic to zoom in on any particular student behaviour and decode its link to assessment. This primarily takes place through students' comments and facial expressions, however from time to time students' actions such as cheating or falling asleep in class highlight particular problems with the assessment. To assist the academic in decoding the students' action, a mentor is on hand to provide an interpretation of the students' responses, illustrated in Figure 2.

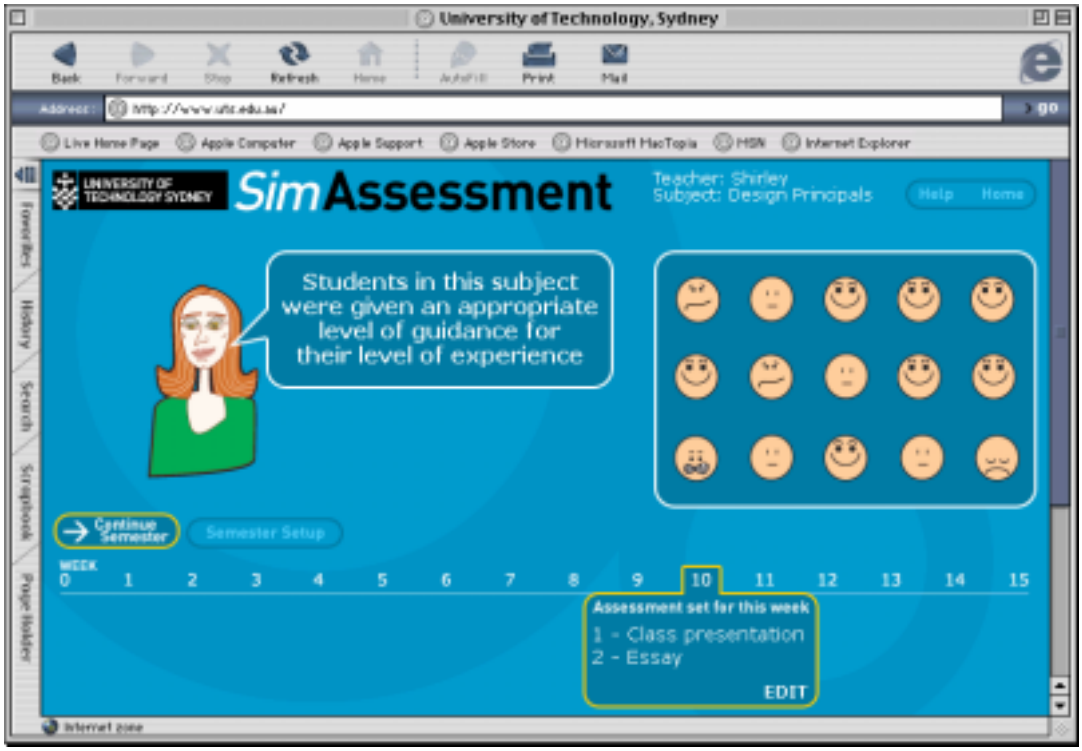

Figure 2. Mentor's response to assessment selection

The debriefing session

What is hoped will make the assessment simulator an effective tool for academic learning is its replication of an experiential learning debriefing session. In the Kolbian model of experiential learning, concrete experience is followed by reflective observation, abstract conceptualisation and active experimentation (Kolb, 1984), The debriefing of the semester's results provides an opportunity for the academic to justify and explain their decisions. The academic is able to interrogate the results of their decisions, formulate hypothesis on why students responded in a particular way and ask the simulator to map alternative responses.

The assessment simulator uses the debriefing session as an area where academics can reflect on the students' actions and experiment with different modes of assessment. To assist in the reflection the academic is joined in the debriefing session by the other four characters in the simulation -- the subject coordinator, their colleague, the tutor and the academic developer. Users can also survey the students to gain their impressions of different assessment tasks.

The rules of the game

At the login session the academic is provided with hints on how to succeed in the simulator. The academic is informed that the goal of this simulator is to get as many students in the class to adopt as deep an approach to their studies as possible. Research has confirmed that the appropriate balance of course goals, student workload and feedback has the greatest chance of achieving these aims (see Ramsden 1992 for details). Further, it is clear that student satisfaction increases with deep approaches to learning, providing better quality outcomes and better grades. Sitting behind the simulation lie a number of assumptions regarding good assessment practices and it is these assumptions that are designed into the responses made by each of the classroom characters.

The basic principle of good assessment embraced by the simulation is that the assessment needs to be aligned to the curriculum goals. Biggs (1999) argues that if the goals of the subject and the assessment tasks are not aligned then the learning outcomes are unlikely to be those the academic had intended. Therefore, the simulator takes goal alignment, calculated by comparing the goals selected by the academic with the reported goals of the faculty, as the foundation for judging the appropriateness of all activities.

To achieve an assessment regime that affords deep approaches to learning the academic needs to have a high level of goal alignment, high levels of feedback and moderate student workload. Student workload also influences the students' emotional state as illustrated in Figure 3. Student workload is calculated using the spacing of assignments, the type of assessment, the weighting and the marker. When the workload is high, the virtual students become increasingly stressed and anxious. Students' actions are less likely to accord with the academics' wishes, and in this simulation, results in high levels of student conflict and cheating. Excessively low workloads lead to lethargy and depression.

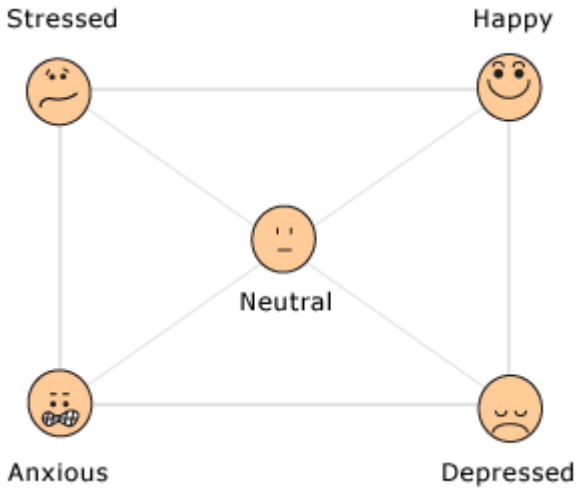

Figure 3. Student character responses to workload

Ramsden's review of the research literature revealed that quality of feedback is often mentioned as the most significant factor in student progress, making provision of effective feedback a difficult but crucial skill (Ramsden, 1992: 99). In this simulator, feedback is calculated on the spacing of the assignments and the type of feedback with the assumption being that academics will have more time to provide this detailed feedback. High levels of feedback result in high quality student outcomes while low levels of feedback involve in the students requiring more information from the lecturer.

Teacher workload is also calculated from the spacing of assignments, the type of feedback, and whether the assessment item will undergo peer-assessment or assessed by the teacher. At the end of the simulation, academics are provided with a summative report that they can use to determine the effectiveness of the choices they made. The six outputs -- goal alignment, approach to learning, student workload, teacher workload, feedback, and public confidence generated by the simulation are presented as qualitative statements similar to those regularly found in student feedback questionnaires. The final report includes an audit of the generic capabilities that may have been assessed as a result of the academic's assessment choices as well as a comparison between the students' approaches to learning at the beginning of the simulator and at the end. We hope the final report will act as a point of comparison for future uses of the simulation. It should also provide an incentive to improve performance and thereby encourage repeat visits to the simulator.

The assessment engine

A critical aspect of the success of any simulation is in the effectiveness of the algorithm developed to model the process being simulated. For this project, the particular challenge was to interpret the choices made by the academic in some way that could be processed by an algorithm. After trialing a number of different approaches, an algorithm based on a weighting system was discussed and eventually selected for use.

The more balanced the assessment pattern, (i.e. a small number of regularly spaced, equally weighted assessment tasks), the higher the algorithm's outcomes. This is somewhat different from the rule-based system (which uses a "if this input", then "that outcome" approach) used in many simulations. Where the input includes a small number of tasks, that would contribute a greater weighting than a large number of

tasks because of the assumption that the student would be able to take a deeper approach to the learning required if the number of tasks was small. Similarly, some weighting was contributed by the timing of the tasks. A well spaced set of assessment tasks would allow students sufficient time to complete each task, and for the teacher to contribute the relevant feedback on their learning. This would attract a higher weighting than, for instance, a choice of 3 tasks each due in week 14 of a semester.

The casting of these choices into a number is an arbitrary task and no significance should be attached to the particular choice of numbers we made. A spreadsheet implementation of the algorithm was developed and tested using a series of reasonable, and extreme, choices. The output from the calculation was then closely inspected for plausibility from a teaching and learning point of view.

Programming

As noted above, the simAssessment prototype was developed on the World Wide Web. A database driven program provided the development team with the flexibility of concentrating on the logic of the programming and enabled us to delay decision making about the content of the different cells to a later time. For example, generic faces and comments have been used for the prototype, but enhancing them at a later time will be a trivial task.

For the backend programming language we chose Macromedia ColdFusion for the prototype because:

- it is suitable for the rapid development of dynamic web pages;
- development time is faster when compared to other similar technologies such as Microsoft ASP;
- it is the main programming language used for web sites developed within IML and therefore programming expertise was available in-house;
- it provides good session and concurrency handling; and finally
- it is able to transparently maintain sessions for multiple users trying to use the system / prototype concurrently.

For the front end we chose Macromedia Flash because of its small file size, as well as the flexibility of files afforded by Flash. These features were required because of the complex nature of the interface of SimAssessment. Secondly, Flash integrates easily with ColdFusion, especially with the latest versions - ColdFusion MX and Flash MX.

Microsoft SQL server was chosen as the backend database for SimAssessment. MSSQL offers good performance and is reliable when handling multiple connections. The database stores all possible values from these assessment parameters, which can be selectively retrieved by ColdFusion in accordance with the user's selection. The database is also used to store all student comments, mentor's comments, and many assessment-specific parameter values. Assessment-specific parameters include many things, such as assessment type, assessment goals, due date, marker, weighting, and feedback type. ColdFusion connects with the database in order to retrieve the values of all assessment parameters which are then used to calculate outcomes.

There were two main technical challenges in developing SimAssessment. First, the process of providing the dynamic outcomes in graphical form in ColdFusion. In order to show the progress of the assessment outcomes dynamically, ColdFusion calculated the outcome values for all assessments up until the current assessment, when the CFX Image custom tag was used to draw the graph.

The second challenge was getting the outcomes to reflect reality. A number of parameters were used to calculate the assessment outcomes which were given a value between 1 and 5. Trial and error was required to determine the value which would give the most realistic outcome.

A number of technical enhancements are already planned for the second version of SimAssessment. The representation of students will be made more complex, for example, the addition of facial expressions such as blinking eyes, and body movements such as raising hands.

Secondly, in the current version, the system only provides users with feedback when the timeline reaches a particular assessment due date. In the future, the system might be designed so that it is also able to provide feedback in between assessment due dates.

Evaluation

SimAssessment has only ever been conceived of as one of a number of strategies for the dissemination of good practice in assessment at UTS. Its specific goals are to experiment with the ways in which an innovative and engaging online activity might provide an alternative means of dissemination of good practice in assessment. The simulator links good practice in assessment with the improvement of teaching and learning as a way of highlighting of the rewards in academics trying out their ideas on assessment. Its focus is on the critical role of assessment in students' experience of their subject and the consequences of this on their approach to learning.

The usability testing of the prototype has been carefully planned and will be conducted during October 2002. The primary purpose of the testing is to understand the strengths and weaknesses of the SimAssessment tool including the identification of usability issues. The specific objectives are to gain an understanding of academics':

- reaction to the overall design and graphic identity, paying particular attention to their reaction to the character representation;
- understanding of the structure of the site and its navigation;
- understanding of the degree to which the project's functionality is useful; and
- reaction to the appropriateness and usefulness of the content.

The evaluation strategies planned include asking test users to complete:

- a pre-questionnaire about their level of experience of confidence in using computers;
- a set of specified tasks; and
- a post-questionnaire to obtain their feedback on the overall value of SimAssessment, its ease-of-use and navigation, usefulness of information, as well as a judgement on whether their colleagues would use a project of this nature.

In developing the final product, the multimedia development process must be started again from the beginning as all aspects of the product design must be revisited to ensure that the usability test findings are included in the product.

- prototype evaluation process and results
- redevelopment
- plans for final product evaluation

Conclusions

This ultimate goal of this project is to improve the quality of learning for students in higher education by providing opportunities for academics to enhance their understanding of good practice in assessment.

The approach is an innovative one, drawing upon Kolb's experiential learning cycle to underpin the development of an online simulation so that academics may test out ideas, receive feedback and reformulate their ideas. This approach has proven to include many technical and design challenges, not the least of which will be academics' acceptance of such an approach.

This is an important experiment in the use of simulation, in that it is critical that we improve students' experience of assessment. As Ramsden (1992: 187) so aptly points out, "From our students' point of view, assessment always defines the actual curriculum…that is where content resides for them, not in lists of topics or objectives".

References

Bamford, A. (2001) Report of the Assessment Project. Internal Document.

Biggs, J. (1999). *Teaching for Quality Learning at University*. Buckingham: Society for Research into Higher Education and Open University Press.

Entwistle, N. (1995) The Use of Research in Student Learning in Quality Assessment. In G. Gibbs (Ed) *Improving Student Learning through Assessment and Evaluation*. The Oxford Centre for Staff Development. pp24 -43.

Entwistle N. J. and Wilson, J. D. (1977). *Degrees of excellence : the academic achievement game.* London : Hodder and Stoughton.

Kolb: D. (1984) *Experiential Learning: Experience as the source of learning and development.* Englewood Cliffs: Prentice Hall.

Pennington, G. (1994). Developing Learning Agents. In P. Nightingale and M. O'Neil (eds) *Achieving Quality Learning in Higher Education*. London: Kogan Page, quoted in Entwistle, N. (1995) The Use of Research in Student Learning in Quality Assessment.

Prosser, M., & Trigwell, K. (1999). Understanding Learning and Teaching: The experience in Higher Education. Buckingham, UK: SRHE & Open University Press.

Ramsden, P. (1992) *Learning to Teach in Higher Education*. London: Kogan Page.

Scott, G. (2002) *UTS Tracking Systems for Teaching and Learning*. Internal Document.

Senate Inquiry in to higher education (2001) *Universities in Crisis*. Available at http://www.aph.gov.au/senate/committee/EET_CTTE/public%20uni/report/contents.htm. Accessed on 28th July, 2001.