Contributed Paper (Reviewed)

Aligning assessment with curriculum and pedagogy in applied information technology

Paul Newhouse, PhD

Edith Cowan University

p.newhouse@ecu.edu.au

ABSTRACT

This paper reports on one part of the first year of a three-year study into the feasibility of using digital technologies to represent the output from practical assessment tasks in senior secondary courses. The aim was to improve the alignment of assessment with pedagogy and curriculum. Students in seven classes studying the *Applied Information Technology* ***(AIT) course in Western Australia completed a digital portfolio and a computer-based exam for assessment. All student work was placed into an online repository to be accessed by external markers through a standard web browser. These were marked using a traditional analytical method and a paired comparisons method. Generally the portfolio was implemented without technical difficulty and for the exam only sound recording provided some difficulty. Both methods of marking provided reliable scores with the paired comparisons method being the more reliable. Students and teachers indicated a preference for assessment at a computer whether to create a portfolio or complete an exam.***

Acknowledgement

The research discussed in this paper is the result of the work of a research team organized by the Centre for Schooling and Learning Technologies at Edith Cowan University (http://csalt.education.ecu.edu.au/). The team was led by Paul Newhouse and John Williams and includes senior researchers Dawn Penney, Cher Ping Lim, Jeremy Pagram, Andrew Jones, Martin Cooper, Alistair Campbell, and many research assistants including David Miller. The work of everyone in this team has contributed to the research outcomes presented in this paper.

INTRODUCTION TO THE PROBLEM

The problem of alignment between assessment, currriculum and pedagogy has been well discussed over the decades, however, this has become more pointed as digital technologies have rapidly changed society and gradually influenced curriculum and pedagogy (Lane, 2004). Many educational researchers argue that traditional assessment fails to assess learning processes and higher-order thinking skills, and go on to explain how digital technologies may address this problem (Lane, 2004; Lin & Dwyer, 2006). Dede (2003) suggests that traditionally educational assessment has been "based on mandating performance without providing appropriate resources, then using a 'drive by' summative test to determine achievement" (p. 6). He goes on to explain how digital technologies may address this problem and claims that "the fundamental barriers to employing these technologies effectively for learning are not technical or economic, but psychological, organizational, political and cultural" (p.9). Kozma (2009) lays out a rationale for change in terms of a mismatch between what is needed in modern society and what is addressed and thus assessed at school. In particular he draws attention to the differences between standardized pen-and-paper assessment and "Tasks in the Outside World".

In Western Australia this problem became critical with the development of new high-stakes senior secondary courses to be implemented over the latter half of the 2000 decade. In most of these courses a major component of student activity involved performance of practical capabilities, in many cases using a variety of technologies, not conducive to assessment using the traditional three-hour paper-based exam. As a result a three-year study was proposed by researchers at the Centre for Schooling and Learning Technologies (CSaLT) at Edith Cowan University (ECU) in collaboration with the Curriculum Council of Western Australia and supported by an Australian Research Council (ARC) Linkage research grant. The study commenced in January 2008, will be completed by December 2010, and concerns the potential to use digital technologies to represent the output from assessment tasks in four senior secondary courses, *Applied Information Technology* (AIT), *Engineering Studies, Italian and Physical Education Studies* (PES). This paper only reports on the first year of the study for the AIT course.

THE PROBLEM WITH AIT

For the AIT course digital technologies provide the content for study as well as pedagogical support. Therefore performance relates to using the technologies to demonstrate capability in using the technologies. The syllabus states that the AIT course "provides opportunities for students to develop knowledge and skills relevant to the use of ICT to meet everyday challenges". As such in the course students should "consider a variety of computer applications for use in their own lives, business and the wider community". In the course students spend the majority of their time in class using digital technologies to develop information solutions. It should therefore be surprising that currently the external assessment consists of a three-hour paper-based exam. This is despite the fact that the

syllabus stipulates that around 50% of the weighting of assessment should be on production.

Unfortunately this discrepancy between the intentions of the course and the method of external assessment was made more serious with the decision that most students were to submit to the external examination process. As a result the 'exam' would have to be appropriate for lower achieving students, it would dominate the course delivery more and would involve a lot more students, increasing the cost considerably. Another confounding change was the requirement for the course to be packaged in a syllabus format with details of specific content for each unit. This led to a shift of focus away from outcomes towards content that immediately highlighted the issue of the variety of relevant contexts that could be involved in the course and the issue of the rapidly changing content of these areas of technology.

While students can include study in AIT towards University entry this would be of limited value if the external assessment propels the course towards becoming mainly 'book work' rather than creative practical digital work. Whether a senior student is aiming to be a mechanic, doctor, accountant or travel agent study in AIT should begin to give them the skills, attitudes and understanding that will support them in being more successful in work and life. Therefore for the study the research problem for the AIT course became that to align with the aims, rationale, outcomes, content and preferred pedagogy assessment must include students using digital technologies. There are a number of ways in which this could be achieved although principally these are either forms of portfolio or of computer-based exam. Among educators there are differences in opinion over which form of assessment would be best with each having strengths and weaknesses. The research question therefore became, which method of assessment, portfolio or computer-based exam or combination, was most feasible for the course?

AN INTERNATIONAL PROBLEM

A comprehensive review of the literature was undertaken, in particular to investigate international research into the use of digital portfolios and computer-based exams. The study connects with two main fields of research: performance assessment, and computer-supported assessment. However, clearly these are subsumed within the general field of assessment.

Computer-Supported Assessment

Computer-Supported Assessment is a broad term encompassing a range of applications from the use of ICT to conduct the whole assessment process (e.g. on-screen testing), to only assisting in one aspect of the task assessment process (eg. recording performance or marking) (Bull & Sharp, 2000). Lin and Dwyer (2006) argue that to date computer technology has only been used substantially in assessment to automate routine procedures such as for multiple-choice tests and collating marks. They suggest that the focus should be on capturing "more complex performances" (p.29) that assess a learner's higher-order skills (decision-making, reflection, reasoning and problem solving) and cite examples such as the use of simulations but suggest that this is seldom done due to "technical complexity and logistical problems" (p.28).

Recently the Joint Research Centre for the European Commission (Scheuermann & Bojornsson, 2009) brought out a major report titled, *The Transition to Computer-Based Assessment*. In it Kozma (2009), while not considering assessment reform to only require the use of ICT, outlines a number of significant advantages in doing so including: reduced costs; increased adaptability to individuals; opportunity to collect process data on student performance; the provision to tools integral to modern practice; and better feedback data. He does introduce a number of challenges to using ICT to support assessment including: start-up costs for systems; the need to choose between standardized and 'native' applications; the need to integrate applications and systems; the need to choose between 'stand-alone' and online implementation; the need for security of data; the need for tools to make the design of tasks easy and efficient; and the lack of knowledge and examples of high-quality assessments supported by ICT. He also highlights methodological challenges including: the extent of equivalence with pen-and- paper; the design of appropriate complex tasks; making efficient and reliable high-level professional judgements; scoring students' processes and strategies; and distinguishing individual contributions to collaborative work.

Assessment of Practical Performance

The call to investigate "performance-and-product assessment" is not new as pointed out by Messick (1994, p. 14) but Lane (2004) claims that a decline in the use of performance assessments has led to a lack of alignment between assessment, curriculum standards, and instructional practices; particularly with regard to eliciting complex cognitive thinking. While Messick does not specifically address digital forms of performance assessment, his arguments for the need to address "issues of validity, reliability, comparability and fairness" apply, particularly to a range of validity criteria. For example, he outlines situations under which product assessment should be considered rather than performance assessment. The issue is their relationship to replicability and generalisability requirements because these are important when performance is the "vehicle" of assessment. At the same time Pollitt (2004) argues that current methods of marking that focus on summing scores on "micro-judgements" is unlikely to accurately measure a student's "performance or ability" (p. 5), and more holistic judgements of performance are required. This body of literature clearly presents the assessment of student performance as critically important but fundamentally difficult with many unanswered questions requiring research.

Two digital forms of assessment were considered for the AIT course, digital portfolios and computer-based exams.

Koretz (1998), who defines portfolio assessment as the evaluation of performance by means of a cumulative collection of student work, analysed the outcomes of four large-scale portfolio assessment systems in USA school systems, in particular, in terms of their reliability. Each example involved marking student portfolios for the purpose of comparing students and/or schools across a state, mainly in English and Mathematics. All of the examples occurred in the 1990s and none involved digital representations of performance. Even though he noted that significant improvements in the implementation and reliable marking of portfolios had been achieved, at that time he saw portfolio-based assessment as "problematic" (p.309). Findings such as this provide a rationale for considering digital solutions to performance assessment. Beetham (2005) provides a list of "issues relating to the use of e-portfolios for summative

assessment" (p. 5). Seven of the nine issues are technical and most are addressed by the use of a good assessment management system.

Computer-based exams involve students sitting at computer workstations completing tasks, including typing answers to questions. They may be required to use various pieces of software to create digital products or may simply use a browser to complete response type assessment. In AIT while both types of assessment activities may be involved it is likely that the focus would be on the former. There has been increased international interest in computer-based testing to assess ICT capability. For example, there have been trials of such tests in a number of countries including the UK, Norway, Denmark, USA and Australia (MCEETYA., 2007). In Australia the ACER used a computer-based test to assess the ICT literacy of Year 6 and 10 students. They developed the test around a simulated ICT environment and implemented the test using sets of networked laptop computers. While they successfully implemented the test with over 7000 students this was over a long period of time and involved an expensive simulated environment that would be difficult to scale for state-wide high-stakes assessment purposes. The trial in the UK also involved a multi-million pound simulated system but was accessed by students through their school computers. In the Norwegian example students used their own government-provided notebook computers. In the USA a decision has been made to include an ICT literacy test in national testing in 2012 but in a number of states there are already such tests.

THE STUDY

The problem being addressed was the need to provide students with assessment opportunities in the AIT course that were on the one hand authentic, where many outcomes do not lend themselves to being assessed using pen and paper over a three hour period, while on the other hand being able to be reliably and manageably assessed by external examiners, without a huge increase in cost.

The first year of the study in 2008 was a 'proof of concept' to explore the feasibility of particular digitally-based forms of assessment. Feasibility was investigated within a framework, developed by Kimbell and Wheeler (2005), consisting of the four dimensions: technological, pedagogic, manageability, and functionality. The study was evaluative in nature set within an ethnographic framework using interpretive techniques involving the collection of both qualitative and quantitative data. It involved seven AIT classes of senior-secondary students, with a total of 115 students participating. A student questionnaire, teacher questionnaire/interview, and student interview proforma were developed.

The AIT Assessment Task

A research team was responsible for developing the assessment tasks. The exact assessment tasks and details related to their development are provided in Appendices in the main report (Centre for Schooling and Learning Technologies, 2009).

It was important that assessment tasks constitute good professional practice, meet the requirements of the course and are reasonably able to be implemented by a 'good' teacher in a real school. A situation analysis was conducted resulting in the identification of content and outcomes conducive to digital forms of assessment for the course. This formed the basis for the development of the assessment tasks. At this point teachers were recruited on the basis that they would agree to implement the assessment tasks within their programme. They were then involved in refining the assessment task(s) and completing the development of the marking criteria and marking key. For the first year a hybrid assessment task structure was developed in order to compare the operation of a portfolio with a performance exam. Details of the five components are provided in Figure 1.

Figure 1: Structure and descriptions of assessment components for AIT in 2008.

	Component	Description
Portfolio	1: Portfolio Digital Product	Students created a digital product prototype of an information solution using applications commonly used in organizations for productivity, planning and communication. The product must have been produced at school over 15 hours and be no more than 20MB in size.
	2: Portfolio Process Document	Students collated a process document related to their digital product over a period of five hours with a maximum of nine pages that comprised four sections: Research, Design, Production and Evaluation.
	3: Portfolio Two Extra Artefacts	Students provided two digital artefacts that illustrated skills in applying design principles in any two domains (e.g. graphics, databases, spreadsheets, web-publishing). The digital artefacts must have been created at school, under supervision from the teacher.
Computer-Based Exam	4: Exam Reflective Questions	A set of reflective questions concerning the digital portfolio to be marked in combination with the portfolio. Students were permitted to view their portfolio product while responding to the questions.
	5: Exam Performance Tasks	A set of six tasks was provided as a scaffold to responding to a design brief. The tasks involved producing a brochure for a holiday resort that included creating a logo, graphs and tables.

AIT Assessment Task Implementation

The AIT assessment tasks were implemented with seven classes (each at a different school) of Year 11 or 12 students studying the AIT Unit 2B within their second semester programme. Although there were some differences in the manner in which the assessment task was implemented for the classes in most ways they were similar. Each class was visited at least four times by a researcher with all sessions conducted by the teacher in a computer laboratory at the school. Each teacher was permitted to set their own design brief for the portfolio product although four used the example, The Miss Shoppe website, provided with the project documentation. All work was supposed to be completed in class but teachers varied in the extent to which this was policed. Typically student portfolio work was submitted by the teacher on a disc and organised by student folders and the computer-based exam files were collected from individual USB flash-drives. For the exam session(s) at least one researcher assisted the teacher in invigilation and provided a USB flash-drive, and a microphone/headphones set. For the exam all student digital work was saved to the flash-drive and typically a copy was also saved to the school's server. Student design work that was done on paper was collected and either scanned or photographed to add to their digital work.

All digital work was transferred to a University server into folders, named using the students' IDs, and with folders and files named in a consistent fashion. The portfolio product was within a folder and the process document was a PDF file. Within each folder there was an index.htm page that was used by the marking tools to display the contents of the folder, this page contained links to the other files. The first artefact folder contained a PDF file of the student's descriptions of the two artefacts. The exam folder contained all the files copied from the exam USB flash drives and PDF version of their plans, brochure, and reflections.

Two online marking tools were developed using the FileMaker relational database software, one for analytical marking and one for paired comparisons marking. The analytical marking tool was designed with the assessment criteria displayed on the left side and the student work on the right. There was a 'Student Results List' screen and five marking screens, each with different criteria and student work. Marks were recorded by clicking on buttons and notes could be made when required. The tool was designed to do all the clerical functions, such as totalling the marks. The paired comparisons marking tool displayed two students' work side-by-side, with the recording of the marker's choices located between them. The marker was required to make four choices, one holistic and three criteria choices, by clicking on large green arrows pointing to the student they wanted to select, and then click on a button to go to the next pair. A field was added for each student so markers could record their notes on the students' work and would not have to view all of the student work each time. The screen shot displayed in Figure 2 shows the layout of the main screen for this marking tool.

Results From First Year

Seven case studies were constructed, one for each class. A full report on each case study can be found in the first main report (Centre for Schooling and Learning Technologies, 2009). The data from all case studies were also combined for analysis as a sample.

Surveys, Interviews and Observations

Data from observations of students completing the assessment tasks were collated to look for consistencies and variations across the cases. Student interviews were conducted after the exam component with at least one small group of students for each case. Teacher interviews were conducted either using email, phone or face-to-face. The external markers were informally interviewed with notes made to identify key points.

Figure 2: Main screen for the paired comparisons marking tool.

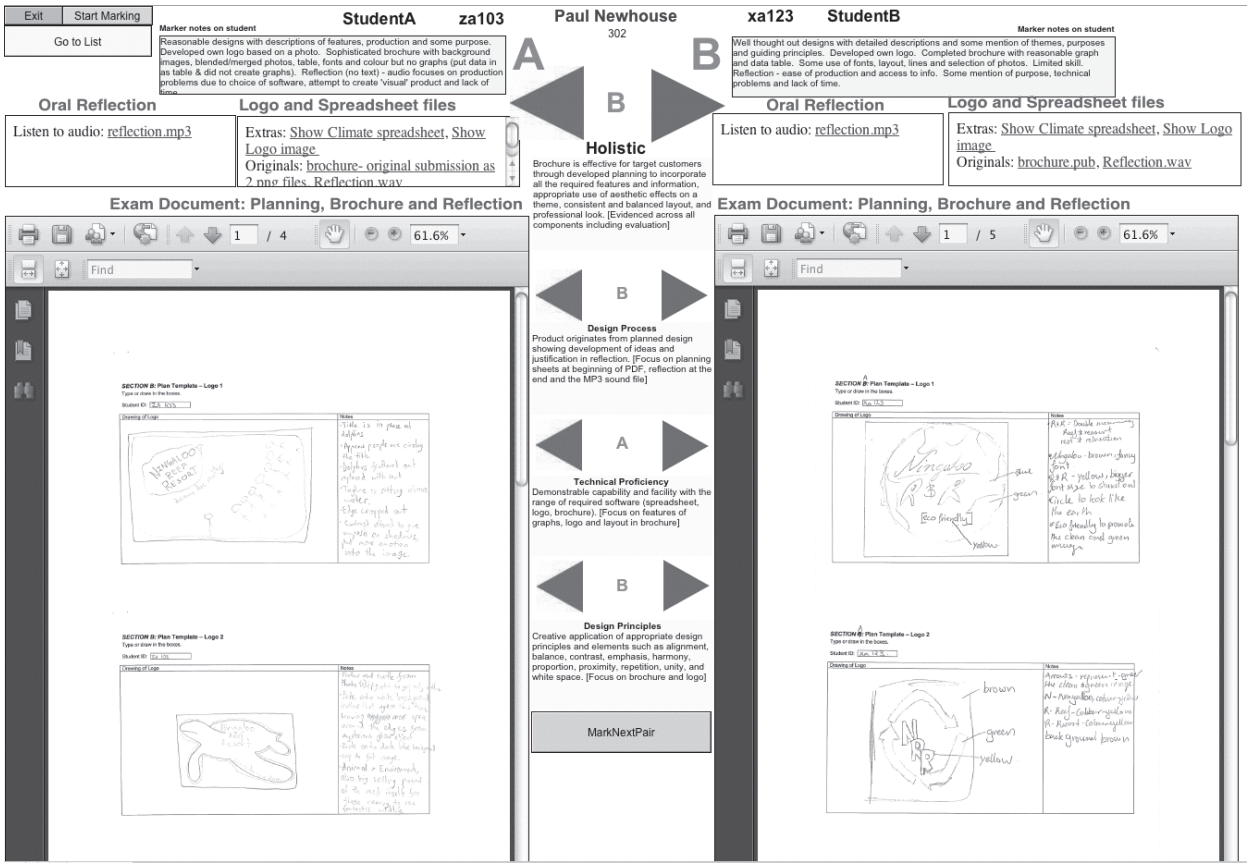
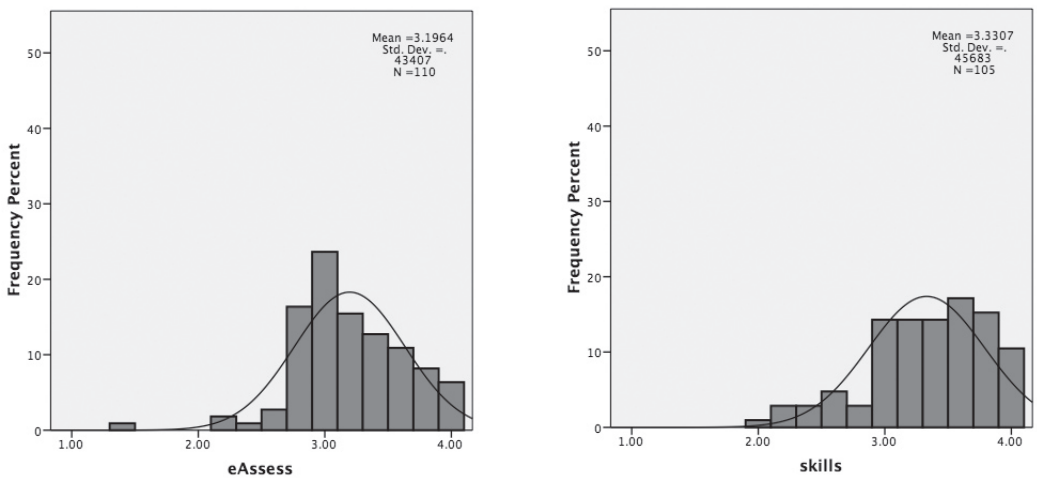


Table 1: Descriptions and descriptive statistics for the scales based on items from the student questionnaire.

	N	Min	Max	Mean	SD	Description
eAssess	110	1.4	4.0	3.2	0.4	Ease of completion of the exam. Score between 1 and 4
eAssessP	108	1.2	4.0	3.2	0.4	Ease of completion of the portfolio. Score between 1 and 4.
Apply	105	1.4	3.0	2.4	0.4	Application of computer use. Score between 1 and 3.
Attitude	105	1.4	3.0	2.6	0.3	Attitude towards using computers. Score between 1 and 3.
Confidence	105	1.0	3.0	2.7	0.4	Confidence in using computers. Score between 1 and 3.
Skills	105	1.0	4.0	3.3	0.5	Self assessment of ICT skills. Score between 1 and 4.
SCUse	105	0.0	334	95.9	62.1	Estimate of time in mins/day using computers at school.

Figure 3: Graphs for the distribution of scores for two scales on the student questionnaire.



For each case the survey of students was conducted immediately on completion of the performance examination. Broadly, it sought students' opinions on the examination and the portfolio, use of computers and other digital devices, attitudes to using computers and facility with computer applications. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from some of the closed-response items. Results for the sample are shown in the Table 1 and two example graphs in Figure 3 on page 8.

Though many students indicated that they had little experience in doing examinations on computers (42% indicated no experience) about half felt they would need little time to get used to the process. Almost all students indicated that doing the examination on the computer was quick, easy and preferable to the traditional pen and paper examination. In general it could be said they were comfortable and positive about both the Portfolio and Exam.

Most students had home access to the technologies listed in the questionnaire with two thirds of the students owning their own laptop computer and 95% having a broadband Internet connection. At school they used computers for an average of 95 minutes per day. All students felt confident with computers and liked using them. Of the types of computer software listed students felt least confident about web authoring and databases. Overall they indicated a high self-assessment of their computer skills with a mean of 3.3 on a four-point scale (in comparison this mean was 2.9 for Physical Education Studies and Italian, and 3.1 for Engineering).

The intention of the practical test was clearly understood namely to design and create a logo for a business and incorporate that into a tri-fold advertising brochure. Only five students did not use a software template for the brochure and of these four used a word-processed document with three columns. Students' responses to questions about the reflective questions section of the exam suggested that either the intention was unclear or that they did not know how to answer. There was widespread confusion over the stages of the technology process and the distinction between these with many responses repeated. As one student noted, "it just seemed like you asked the same questions four times… I got four words out of a thesaurus and copied and pasted those in three or four times".

In general the teachers were very positive towards both forms of assessment and felt that these complemented their own aims, principles and methods of instruction. All teachers said they would like to see a greater emphasis on the practical aspects of the course. Some were cautious about the potential of the examination because of the possibility of technical problems.

Analysis of Results of Marking

Two external assessors (experienced computing teachers) marked each student's work using the analytical marking tool. Five assessors, including these two, also used the paired comparisons marking tool to mark the practical examination for a reduced population of 60 students. These students were chosen because their practical work samples were equivalent in the degree of completeness and had no missing sections. In particular they all had an audio response file for the exam. These assessors completed a pre-determined set of comparisons. Of the 115 students, only 58 final semester and 26 assessment task marks were received from teachers.

The amount of time taken in the analytical marking varied from an average of about 5 minutes to 25 minutes per student. The shorter times corresponded to students whose submissions of the five components were incomplete. The longer times were associated with student work which comprised large files (for example some animations were several MB) and these took time to download to the assessor's computer prior to marking. Apart from delays in downloading, longer time was spent on higher quality work, particularly where evidence of performance had to be gathered from several parts of the submission. For the comparative pairs marking the time required to make a comparison was initially just under 10 minutes, and overall about 3 minutes per comparison.

Analytical Marking

Both external assessors gave a very similar range of scores with a resulting similar mean and standard deviation. There was no significant difference between their means on the overall scoring of the assessment task (Portfolio and Exam combined). The range of scores, the means and standard deviations varied considerably between individual classes on separate components of the assessment task and for the task as a whole. For example, one class had a mean of 53.6 while another had a mean of 31.2.

There was a strong and significant correlation between the two markers with correlation coefficients of 0.89 ($p<0.01$) for the scores and 0.91 ($p<0.01$) for the ranking of students. Correlation between the average mark of the external assessors and the semester mark awarded by the teacher was moderately strong and significant ($r=0.62$, $p<0.01$). There was no significant correlation between the average mark of the external assessors and the marks awarded by the teachers.

The results of marking of the portfolio and exam components was analysed separately. Summary statistics are shown in Table 2 for the external markers (average of the two). The mean 'external' score for the exam was around 50% but for the portfolio only around 37%. This was because many students did not submit all components of the portfolio whereas almost all students completed exam. There was a moderate but significant correlation between marks awarded by the external assessors for the examination and marks awarded for the portfolio ($r=0.58$, $p<0.01$). An Analysis of Variance was completed to investigate variance across the classes of students. As expected the marks awarded by the external markers varied significantly ($p<0.01$) by class for the Portfolio and Exam (and components of each) indicating differences in the capability of students between classes. The differences were consistent with the same classes having higher means for both the Portfolio and Exam.

A Rasch model was applied to the Exam and Portfolio marks using the responses of both markers to generate a combined score for each student. For the exam marks there was one reversed threshold and as a result the analysis was repeated with two responses combined. This increased the SI marginally to 0.853 and removed the reversed threshold. There were few extreme outliers with a frequency distribution relatively well spread. The analysis also gave a reliable set of scores for all three components of the portfolio (SI=0.96, 0.96 and 0.94 respectively). There were a few extreme outliers that tended to be students scoring 0 on all or almost all of the criteria. The frequency distributions tended to be well spread, with high standard deviations and not very 'normal' in structure. No modifications were required although the thresholds for three of the criteria did not work very well. In general the analytical marking of the portfolio components gave reliable scores when taken separately. However, some improvements in the marking criteria could be made. Using t tests to test for differences in means between the markers found a significant difference for the overall score but only a significant difference ($p < 0.05$) on two of the separate criteria. For both these criteria, at least one potential response was used sparingly.

Comparative Pairs Marking

One holistic and three specific assessment criteria were developed for the comparative pairs marking from the criteria previously developed for the task. The RUMMcc software used to apply the Rasch model to the pairs marking data provided an 'Estimate' score for each student as a location in logits (logarithmic units of measurement) with a standard error of measurement, for each criteria. A Separation Index (SI) (value between 0 and 1) was calculated as an indicator as to whether or not the exemplars were sufficiently diverse in quality to assure a broad enough range for the purposes of comparison. The SI for the Holistic criterion was 0.958 indicating a highly reliable set of scores (values above 0.8 are considered to be good). Intra-rater reliability analysis gave a group reliability was 1.01 where this statistic should be between 0.5 and 1.5.

There was a strong and significant correlation between the three specific criteria and the holistic criteria and between the three

Table 2: Descriptive statistics for analytical marking of exam and portfolio.

Component	N	Possible	Minimum	Maximum	Mean	Std. Dev	Mean (%)
Exam	113	30	6.0	24.5	14.8	4.2	49.3
Portfolio	98	70	4.0	52.5	26.2	12.2	37.4
Total	96	100	13.5	74.5	41.2	15.2	41.2

Table 3: Correlations between assessors, assessors and teachers, and methods of marking for the AIT assessment (N=115).

Analytic Assessors		Assessors and Teacher		Analytic and 'Pairs'	
Marks	Rankings	Marks	Rankings	Marks	Rankings
0.89**	0.91**	0.32	0.30	0.73**	0.71**

** p<0.01 (2-tailed) * p<0.05 (2-tailed)

(r>0.74, p<0.01). There was also strong and significant correlation (r=0.73, p<0.01) between the mark generated by comparative pairs marking and the mark determined by analytical marking. As might be expected, the criteria (pairs marking) are also strongly correlated with the average practical mark (analytical marking). There was no significant correlation between the teacher's examination mark and the pairs marking, with the exception of criterion 2 (r=0.46 p<0.05). However, the teacher's semester mark was weakly correlated with all criteria in the pairs marking.

Conclusions about Marking

For each case study and the entire sample the results of marking were compiled into a table for scores and rankings. The rank for the average of the external assessors marks was for all 115 students whereas the teacher's semester mark rank was just for the class. The ranking from the 'pairs' marking was for the 60 students whose exam was marked in this manner. Correlations between these results are summarised in Table 3.

Which method of marking delivered the most reliable results? Only the practical component of the Exam (Component 5) was marked with both methods of marking. There was a relatively low correlation (r=0.43, p<0.01) between the two external 'analytical' markers although their average was relatively highly correlated to the results of the pairs marking (r=0.73, p<0.01). Rasch analysis showed that the exam scores were highly reliable (SI=0.93) using the comparative pairs method and with a minor modification to one criterion reasonably reliable (SI=0.85) from the analytical marking.

Which form of assessment, portfolio or exam, was marked the most reliably? Only two schools implemented all aspects of the portfolio largely in line with the stated requirements. An analysis of their combined data found that the exam and portfolio scores were significantly correlated (r=0.71, p<0.01). The correlation between the two external analytical markers was significant for the Portfolio (r=0.85, p<0.01) and Exam (r=0.54, p<0.01). This would suggest that the portfolio could be more reliably marked than the exam.

CONCLUSIONS

A summary of findings was compiled based on the feasibility framework and including a summary of the constraints and benefits of the form of assessment. In general it could be concluded that both the portfolio and computer-based exam were implemented successfully although the computer-based exam was the easiest to implement in a consistent fashion between schools. However, the portfolio was more reliably marked using the analytical method. The only implementation issues for the exam were the failure of audio recording in three schools and the handful of students who had to move workstations mid-exam due to technical failure at a cost of no more than five minutes. The collation of student work in digital form had several obvious advantages for assessment such as ease of storage, backup, transmission, access and sharing. The web based database, which held the student work was responsive and easy to use provided that adequate Internet bandwidth was available to the marker.

Overall there was no compelling reason to choose one form of assessment over the other with each having strengths and weaknesses. The Portfolio allowed students to adequately demonstrate their capability with some scope for tailoring to the context for the student. However, ideally it would need an online management system and would need a well-structured system for verification that would probably include some type of signed affidavit with spot checks on a sample of students to ensure all teachers implemented the portfolio according to the required conditions.

For the Exam the study highlighted two areas in which decisions would need to be made: technical implementation; and performance tasks specification. The project used USB flash drives that worked on school computers for all students. This would be cumbersome but not unrealistic to scale up but in the long-term an online exam management system should be used. However, this introduces the school network that was found to be a confounding variable. No attempt was made to limit access to software that was normally available to the students. The low-end nature of the tasks meant that this provided little, if any, advantage to any students with most using Microsoft software and a relatively basic graphics package. A major weakness of the exam was that a relatively low level of competence was required (logo, brochure, and spreadsheet graph) so that students in all classes could attempt them. The types of practical tasks teachers give students varies tremendously (e.g. many do not give database type of tasks or even spreadsheets while others will not do animations or movie editing). Therefore it is very difficult to set tasks that all students would have the background to attempt and would also allow adequate scope for the more capable students. This is not a problem with the portfolio as the design brief could vary between classes and typically allowed tasks to be relatively open-ended. In the second year the study will trial a more complex task for the exam that allows for more choice.

REFERENCES

Beetham, H. (2005). *e-portfolios in post-16 learning in the UK: developments, issues and opportunities.* Bristol, UK: Joint Information Systems Committee (JISC)

Bull, J., & Sharp, D. (2000). *Developments in computer-Assisted Assessment in UK Higher Education.* Paper presented at the Learning to Choose: Choosing to learn, QLD, AUS.

Centre for Schooling and Learning Technologies. (2009). *Investigating the feasibility of using digital representations of work for authentic and reliable performance assessment in senior secondary school courses. First Year Report.* Perth: Edith Cowan University

Dede, C. (2003). No cliche left behind: why education policy is not like the movies. *Educational Technology, 43*(2), 5-10.

Kimbell, R., & Wheeler, T. (2005). *Project e-scape: Phase 1 Report.* London: Technology Education Research Unit, Goldsmiths College

Koretz, D. (1998). Large-scale portfolio assessments in the US: Evidence pertaining to the quality of measurement. *Assessment in Education, 5*(3), 309-334.

Kozma, R. B. (2009). Transforming Education: Assessing and Teaching 21st Century Skills. In F. Scheuermann & J. Bojornsson (Eds.), *The Transition to Computer-Based Assessment* (pp. 13-23). Ispra, Italy: European Commission. Joint Research Centre.

Lane, S. (2004). Validity of High-Stakes Assessment: Are Students Engaged in Complex Thinking? *Educational Measurement, Issues and Practice, 23*(3), 6-14.

Lin, H., & Dwyer, F. (2006). The fingertip effects of computer-based assessment in education. *TechTrends, 50*(6), 27-31.

MCEETYA. (2007). *National Assessment Program – ICT Literacy Years 6 & 10 Report 2005.* Carlton South, Australia: Curriculum Corporation

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.

Pollitt, A. (2004, June 2004). *Let's stop marking exams.* Paper presented at the International Association for Educational Assessment Conference, Philadelphia.

Scheuermann, F., & Bojornsson, J. (Eds.). (2009). *The Transition to Computer-Based Assessment.* Ispra, Italy: European Commission. Joint Research Centre.

BIOGRAPHY

Paul Newhouse (Ph.D) *is an associate professor in educational computing at Edith Cowan University in Perth, Western Australia. He is currently the director of the Centre for Schooling and Learning Technologies (CSaLT) in the School of Education. He has always considered himself to be both a teacher and learner from early years in an innovative state secondary school in Western Australia to conducting research in schools and working with pre-service and practicing teachers at two universities. Throughout his aim has been to improve the opportunities for all children to develop as decision-making, empowered, responsible, relational citizens through engaging and relevant schooling. His focus is on implementation strategies for using ICT to support learning in schools, particularly as applied to portable computing, assessment and curriculum development in technology education.*