Digital Forms of Assessment
ECU-Industry Collaborative Research
2006/7
Digitally Based Formats for Alternative External Assessment for Senior Secondary School Courses in W.A.

Final Report

A research project supported by a grant from the ECU Industry Collaborative Scheme

[Project Code: 06-138 NEWHOUSE]

For the Curriculum Council of Western Australia

April 2008

Centre for Schooling and Learning Technologies (CSaLT)

Edith Cowan University
Research conducted by a team from Edith Cowan University under the leadership of Associate Professor Paul Newhouse and Dr. John Williams.

Researchers: Associate Prof. Paul Newhouse  
Dr. John Williams  
Dr. Jeremy Pagram  
Ms. Rosemary Naughton

Manager: Mrs. Lorraine Kershaw  
Research Assistance: Mr. Alistair Campbell, Dr. Martin Cooper

© 2008 School of Education, Edith Cowan University, Western Australia.

This is the final report to the Curriculum Council of Western Australia of research exploring the potential of using digital technologies to support assessment tasks for summative high-stakes purposes in the Applied Information Technology and Engineering Studies courses of study. This report provides an analysis of all the data collected from the sample of schools from October 2006 to November 2007.

NOTE: This report is only for the Curriculum Council of Western Australia and should not be quoted in any context or given to any other person without the express permission of the authors. Members of the research team at Edith Cowan University have copies and a hard copy is kept securely by the Office of Research Innovation.
Contents

EXECUTIVE SUMMARY ................................................................................................................. 4

1 - INTRODUCTION AND BACKGROUND .................................................................................. 7
   1.1 SIGNIFICANCE AND RATIONALE ...................................................................................... 7
   1.2 STATEMENT OF PROBLEM AND RESEARCH QUESTION ................................................. 9

2 - LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK .................................................. 10
   2.1 COMPUTER-SUPPORTED ASSESSMENT ............................................................................ 10
   2.2 ASSESSMENT OF PRACTICAL PERFORMANCE ............................................................... 11
   2.3 DIGITAL FORMS OF ASSESSMENT .................................................................................. 11
   2.4 METHODS OF MARKING ................................................................................................. 13
   2.5 DIGITAL REPRESENTATIONS OF ASSESSMENT TASKS .................................................. 14
   2.6 GUIDELINES FOR DESIGN AND DEVELOPMENT OF DIGITAL ASSESSMENTS .............. 15
   2.7 CONCEPTUAL FRAMEWORK FOR THE STUDY ............................................................. 18

3 - METHOD ................................................................................................................................... 20
   3.1 SAMPLES .......................................................................................................................... 21
   3.2 DATA COLLECTION AND ANALYSIS ............................................................................... 21
   3.3 METHODOLOGY FRAMEWORK ....................................................................................... 23

4 - APPLIED INFORMATION TECHNOLOGY CASE STUDIES .................................................. 26
   WA: WEBSITE PORTFOLIO BASED ON TEMPLATE ............................................................... 27
   AA: DIGITAL PORTFOLIO SLIDESHOW .................................................................................. 33
   IA: ELECTRONIC PORTFOLIO USING FLASH ..................................................................... 40
   PA: DIGITAL PORTFOLIO WITH FORMAT CHOICE .............................................................. 48
   MA: VIDEO PRODUCTION EXAM ......................................................................................... 54
   CA: WEB-PAGE PRODUCTION EXAM ................................................................................. 63
   XA: POSTER GRAPHICS TASK EXAM ................................................................................. 70

5 - ENGINEERING CASE STUDIES ............................................................................................. 80
   GE: RESEARCH PORTFOLIO ..................................................................................................... 81
   ME: PORTFOLIO BASED ON TEMPLATE ............................................................................. 85
   VE: FOCUSED PERFORMANCE TASK ................................................................................... 90
   WE: PORTFOLIO BASED ON RESEARCH .......................................................................... 95
   HE: EXTENDED PRODUCTION EXAM ................................................................................ 99
   YE: FOCUSED PERFORMANCE TASK ............................................................................... 104

6 - FINDINGS, CONCLUSIONS AND RECOMMENDATIONS ..................................................... 109
   6.1 CONCLUSIONS FROM AIT CASE STUDIES .................................................................... 110
   6.2 SUMMARY OF FINDINGS FROM AIT CASE STUDIES .................................................. 112
   6.3 CONCLUSIONS FROM ENGINEERING CASE STUDIES ................................................ 115
   6.4 SUMMARY OF FINDINGS FROM ENGINEERING STUDIES CASE STUDIES .................. 117
   6.5 OVERALL CONCLUSIONS .............................................................................................. 119
   6.6 CONCLUSIONS AND RECOMMENDATIONS .................................................................. 123

REFERENCES ................................................................................................................................. 126

APPENDIX A: DESIGN BRIEF FOR APPLIED INFORMATION TECHNOLOGY ............................. 128
APPENDIX B: DESIGN BRIEF FOR ENGINEERING STUDIES ................................................ 139
APPENDIX C: STANDARDS-BASED MARKING TOOL ................................................................... 141
APPENDIX D: COMPARATIVE PAIRS MARKING TOOL ............................................................... 142
Executive Summary

This report is the result of a small exploratory study conducted 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 ECU-Industry collaborative research grant. The study commenced in August 2006, was completed by December 2007, and concerned the potential to use digital technologies to represent the output from assessment tasks in two senior secondary courses, Applied Information Technology (AIT) and Engineering Studies (Eng or Engineering).

While students tend to focus on, and be motivated by practical performance in courses, teachers being accountable for student results will however tend to ‘teach to the test’. Further, educators are accountable to society for the outcomes of the use of resources in education, and our society increasingly expects that students should demonstrate practical performance not just theoretical knowledge. Moreover, students are more likely to experience deep learning through complex performance. As McGaw (2006) explains, this places a responsibility on education authorities to consider strategies to increase the assessment of performance on practical tasks.

“If tests designed to measure key learning in schools ignore some key areas because they are harder to measure and attention to those areas by teachers and schools is then reduced, then those responsible for the tests bear some responsibility for that.” (McGaw, 2006, p. 3)

Performance-based assessment is not new. Oral and laboratory examinations have been used in European schools and Universities for over a century. In many industries performance-based assessment approaches are used (e.g. pilots). In many high-stakes courses in developed countries performance is, and has been, assessed using observation, interview, portfolio or recording (e.g. USA, UK, Denmark). For example, a recent review of assessment methods in medical education (Norcini & McKinley, 2007) outlines performance-based assessment of clinical, communications and professional skills using observations, recordings and computer-based simulations. In W.A. there has been a history of performance-based assessment in some courses in the Arts. However, the use of performance-based assessment in high-stakes courses has been limited by the costs involved in collecting the evidence of performance and difficulties in ensuring reliable and valid results.

Recent advances in psychometric methods and improvements in digital technologies provide tools to assess a variety of performance relatively cost-effectively. However, any approach or strategy will not be perfect and will require compromises and consideration of the following questions.

1. What skills or knowledge are best demonstrated through practical performance?
2. What are the critical components of that practical performance?
3. Why can’t those components be demonstrated on paper?
4. What alternative representations other than paper could be used?
5. What level of compromise in reliability, authentication and cost is acceptable in preference to not assessing the performance at all?

Most of the new courses in senior secondary school in Western Australia have some practical components and there is an expectation from students and community that the assessment of student performance will reflect the nature of this learning. In Western Australia, for most courses, students are externally assessed using traditional methods employing predominantly paper and pen technologies. In most cases it is clear that performance on practical tasks cannot be assessed adequately using paper and pen. Many educational researchers argue that traditional assessment only measures knowledge of basic facts and procedures but fails to assess learning processes and higher-order thinking (decision-making, reflection, reasoning and problem solving).

The focus of this study was on the use of digital technologies to ‘capture’ performance on practical tasks for the purpose of high stakes summative assessment. The purpose was to explore this potential so that such performances could be included to a greater extent in the assessment of senior secondary courses to increase the authenticity of the assessment in these courses. The study involved seven small
case studies for AIT involving eight teachers and eight classes, and six small case studies for Engineering involving six teachers and seven classes. The number of students involved in each case study ranged from nine to 48. While caution needs to be taken in interpreting the analysis and generalising from the results the study has provided the opportunity to develop a sound methodology, including analysis techniques, and has suggested areas for further research and some guidelines for policy and practice.

Three different fundamental forms of assessment (reflective portfolios, production exams, and performance tasks exams) were investigated in 13 cases with students from either the AIT or Engineering courses and with the assessment task being different in each case. For each case a variety of quantitative and qualitative data was collected from the students and teachers involved including digital representations of the students’ work on the assessment tasks. These data were analysed and used to address the research questions within a feasibility framework consisting of four dimensions: Manageability (Can the digital assessment piece be reasonably managed in a normal classroom?), Technical (Can existing technologies be adapted for digital assessment purposes?), Functional (Is the digital assessment data reliable and valid when compared to traditional forms of assessment?), and Pedagogic (does a digital form of assessment support and enrich students’ learning experiences?).

Within this feasibility framework a range of digital assessment types were employed allowing the researchers to explore a number of different scenarios. Amongst these assessment types were production exams, performance task exams, portfolios, and reflective portfolios. In all there were 13 different assessment tasks across the two subjects each implemented uniquely in a school. Typically these digital assessment tasks were marked by two external assessors and independently by the classroom teacher, almost always using detailed sets of criteria that were represented as a rubric and linked to the appropriate course outcomes. Correlations were determined for comparison purposes between the two external assessors and also between the assessors and the classroom teacher. Additionally, one of the AIT assessment pieces was marked using the method of comparative pairs and these results were again compared to other forms of marking.

Results

The study involved seven AIT teachers and seven Eng teachers and their senior-secondary students with student numbers varying from 9 to 48 for a given task. The AIT students worked on portfolios, which were presented in Flash, Powerpoint, or as a web page, production exams where they made a web page or a video, or a performance tasks exam where they had to demonstrate graphics skills to make a poster. The Engineering students worked on design portfolios, performance exams, and production exams across a number of different contexts.

Findings are summarised below under the headings for the feasibility framework.

Manageability

An important finding of the study was that in all cases there was sufficient technical knowledge, skill, and IT infrastructure stability to ensure the assessment tasks could be completed to a satisfactory level. Teachers were able to competently implement the assessment tasks although occasionally a lack of student skills and technical knowledge limited the value of the assessment tasks. There were some logistical difficulties when managing longer period exams or portfolios. These difficulties usually manifested themselves in terms of students not having adequate time to complete all requirements. Generally, portfolios were easier to manage when they had more structured templates and enough time was allocated.

Occasionally there were difficulties for assessors accessing student work particularly where there was a large range of possible digital formats. Therefore, consideration of file formats and the size of files was vital to the success of the assessment tasks. It was found to be most efficient to limit file formats to html, pdf, jpg, gif, or avi.
Technical

Overall there were not any significant technical difficulties. Student work was usually collected on a DVD or USB drive and this process worked well. Some maintenance was required in both to prepare completed student work for the assessors. The formats that required least work were pdf, web, or avi formats. Large file sizes particularly for video formats posed issues when students’ work needed to be saved and passed on to the markers.

Functional

In almost all cases the students perceived the assessment tasks to be authentic and meaningful and preferred the task to a written exam. In many of the cases though a lack of time limited the opportunity for students to demonstrate higher levels of achievement.

In terms of correlation between the external assessors the study concluded that this correlation is improved when the assessment rubric contains information specific to the task but still based on the standards framework for the course. Further, to ensure highly reliable marking, consensus discussions are required between markers. Correlations between the external markers and the teacher were uniformly low. This may be because when assessing a task teachers take into account background knowledge about the students and/or background knowledge about the task. This may imply that teachers should not mark their own students output.

The comparative pairs marking undertaken for one of the assessment items produced highly reliable scores. However, questions remain concerning the validity or results with variations for some students compared to the standards-based assessment.

Pedagogic

A further key finding was that there was a fine balance between providing adequate scaffolding to the task and allowing students to demonstrate understanding by making independent choices. Greater structure assisted the marking process but this came at the risk of stifling student creativity and decision making. Clearly the design of the task and the instruction set provided to students are critical to the overall success of an assessment item. Assessment tasks were also more successful when the approach was familiar to the students. For reflective portfolios this was most often the case whereas for exams this was more dependent on the design of the task. In general, in this study the production and performance exams used approaches that were familiar to students and consistent with the pedagogy of the courses.

Constraints and Conclusions

A major constraint that was identified through this study was the amount of time the students required to complete the task. This was particularly the case when students were constructing reflective portfolios. Additionally, and mainly in AIT, a lack of student ICT skills, organisation, and literacy limited students ability to demonstrate their understandings.

The overall finding was that the benefits outweighed the constraints. In both courses students and teachers were generally positive about the assessment tasks and their implementation. The students overwhelmingly preferred more practical work, which was of greater relevance and interest, to pen and paper exams. Most importantly, in all cases the assessment tasks could be completed at school with school equipment and the teacher could authenticate the student output.
1 - Introduction and Background

This report has resulted from a study carried out from August 2006 to December 2007 by researchers from the School of Education at Edith Cowan University (ECU) and managed by the Centre for Schooling and Learning Technologies (CSaLT). The study set out to investigate the use of digital forms of assessment in two upper secondary school courses. The study was done in collaboration with the Curriculum Council of WA and supported by an ECU-Industry collaborative research grant.

The study built on concerns that the assessment of student achievement should, in many areas of the curriculum, include practical performance and that this will only occur in high-stakes context if the assessment can be shown to validly and reliably measure the performance and be manageable in terms of cost and school environment. The assessment is summative in nature (i.e. it is principally designed to determine the achievement of a student at the end of a learning sequence rather than inform the planning of that sequence for the student) with reliability referring to the extent to which results are repeatable, and validity referring to the extent to which the results measure the targeted learning outcomes.

The study specifically addressed a critical problem for the school systems in Western Australia, which also has national and international significance. At the same time the study advanced the knowledge base concerning the assessment of practical performance by developing techniques to represent practical performance in digital forms, collate these in online repositories, and judge their quality using a standards-based marking method and trialing a comparative pairs marking method.

The study was designed to be a one-year pilot study conducted with the Curriculum Council with a view to developing a larger more long-term study. It explored the use of three digital forms of assessment in two courses with a small number of teachers and their classes.

1.1 Significance and Rationale

From the 1990s, significant developments in computer technology have been the emergence of low-cost, high-powered portable computers, and improvements in the capabilities and operation of computer networks (e.g. intranets and the accessibility of the Internet). These technologies have appeared in schools at an escalating rate. During that same period school systems in Australia were moving towards a more standards-based curriculum and investigating methods of efficiently and effectively assessing students from this perspective. In Western Australia this became critical with the development of high-stakes senior secondary courses to be implemented over the latter half of the decade. In some courses developments in technology dictated that students should be assessed making use of that technology while in many courses it was likely that at least some of the intended learning outcomes were not able to be adequately assessed using paper-based methods. Therefore it was important that a range of forms of assessment were considered along with the potential for digital technologies to support these alternative forms.

There is a critical need for research into the use of digital forms of representation of student performance on complex tasks for the purposes of summative assessment that are feasible within the constraints of school contexts. This study investigated authentic digital forms of assessment with high levels of reliability, manageability, which were capable of being scaled-up for statewide implementation in a cost effective manner. The findings of this study provide guidelines for educational administrators that reflect successful practice in using Information and Communications Technology (ICT) to support a standards-based approach to assessment. The findings of this study provide significant benefit to the partner institution and the wider educational community, particularly in terms of the national priority afforded the provision of a nationally consistent schooling system with greater accountability to standards in senior schooling across Australia. This provides a national imperative to develop rigorous, reliable, and viable forms of performance assessment.

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). This argument centres around the validity of the assessment in terms of the intended learning outcomes, where there is a need to improve the criterion-related validity, construct validity and consequential validity of high-stakes assessment (McGaw, 2006).
Further, in some school courses students learn with technologies and this dictates that students should be assessed making use of those technologies. 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).

Lin and Dwyer (2006) argue that to date computer technology has really 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 and the SMART model but suggest that this is seldom done due to "technical complexity and logistical problems" (p.28). The SMART (Special Multimedia Areas for Refining Thinking) model is an online curriculum and assessment tool developed by the Cognition and Technology Group at the Vanderbilt University (Cognition and Technology Group, 1996).

Apart from the lack of validity of traditional paper-based assessment methods another compelling rationale to consider the efficacy of performance assessment is that teachers tend to teach to the summative assessment (Lane, 2004; Ridgway, McCusker, & Pead, 2006). McGaw (2006) discussed this in the light of changes in the needs of the society, advances in psychometric methods, and improvements in digital technologies and believed that this was critical due to the influence of assessment on the judgements of teachers and students concerning the curriculum with “risk that excessive attention will be given to those aspects of the curriculum that are assessed” and that “risk-taking is likely to be suppressed” (p.2). He goes as far as to argue that, “If tests designed to measure key learning in schools ignore some key areas because they are harder to measure, and attention to those areas by teachers and schools is then reduced, then those responsible for the tests bear some responsibility for that” (p. 3). A concern underpinning the argument for computer-based assessment methods to replace traditional paper-and-pencil methods was presented by the American National Academy of Sciences (Garmire & Pearson, 2006). They argue that assessing many performance dimensions is too difficult on paper and too expensive using “hands-on laboratory exercises” (p. 161) while computer-based assessment has the potential to increase “flexibility, authenticity, efficiency, and accuracy” but must be subject to “defensible standards” (p. 162) such as the Standards for Educational and Psychological Testing (AERA et. al., 1999). The committee cites the use of computer-based adaptive testing, simulations, computer-based games, electronic portfolios, and electronic questionnaires as having potential in the assessment of technological literacy (2006). They concluded that computer-based simulations were suitable but could be expensive. They also raised a number of questions requiring research that electronic portfolios, “appear to be excellent tools for documenting and exploring the process of technological design” (p. 170).

McGaw (2006) also believes that without change to the main high-stakes assessment strategies currently employed there is a reduced likelihood that productive use will be made of formative assessment. He is not alone in this concern, for example, Ridgway (2006, p. 39) states that, “There is a danger that considerations of cost and ease of assessment will lead to the introduction of ‘cheap’ assessment systems which prove to be very expensive in terms of the damage they do to students’ educational experiences.” Therefore, from both a consideration of the need to improve the validity of the assessment of student practical performance, and the likely negative impact on teaching (through not adequately assessing this performance) there is a strong rationale for exploring alternative methods of assessment.

While the use of some digitally based forms in assessment have been trialed to some extent in other parts of Australia and overseas, this has not been done in Western Australia, and there has been little comparative research internationally into the variety of digitally based forms that may be considered. Further, it is important that such research be conducted within typical school-based settings to reflect the realities of students. This study sought to conduct an in-depth exploration of the variety of forms of digitally based assessments in a number of different courses with the view to expand this to be applicable to a range of courses.
The findings of this study provide significant benefit to the partner institution and the wider educational community, and guide educational administrators in the effective implementation of ICT at a system level. The study sought to develop a set of guidelines that reflect successful practice in using ICT to support a standards-based approach to education.

1.2 Statement of Problem and Research Question

The general aim of the study was to explore the potential of various digitally-based forms for external assessment for senior secondary courses in Western Australia. Specifically the study set out to determine the feasibility of three digital-assessment forms in terms of manageability, cost, validity and reliability, and the need to support a standards-based curriculum framework for students in schools across the state. The problem being addressed was the need to provide students with assessment opportunities in new courses, that are on 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. That is, the external assessment for a course needs to accurately and reliably assess the outcomes without a huge increase in the cost of assessment.

The main research question was:

How are digitally based representations of student work output on tasks most effectively used to support highly reliable summative assessments of student performances for courses with a substantial practical component?

This study addressed this question by considering a number of subsidiary questions.

1. What are the benefits and constraints of each digitally-based form to support the summative assessment of student practical performance in senior secondary courses in typical settings?
2. What is the feasibility of each digital form of assessment in terms of the four dimensions: technical, pedagogic, manageability, and functional?
3. Does the paired comparison judgements method deliver reliable results efficiently when applied to student practical performance?
2 - Literature Review and Conceptual Framework

The study connected with two main fields of research: performance assessment, and computer-supported assessment. However, clearly these are subsumed within the general field of assessment. While it will be assumed that the basic constructs within the field of assessment are known and apply perhaps it is useful to be reminded of this through a statement of the three pillars that Barrett (2005) suggests provide the foundation for every assessment rests.

1. A model of how students represent knowledge and develop competence in a content domain.
2. Tasks or situations that allow one to observe students’ performance.

2.1 Computer-Supported Assessment

Computer-Supported Assessment, sometimes referred to as Computer-Assisted Assessment, is a broad term encompassing a range of applications from the use of computers to conduct the whole assessment process such as with on-screen testing, to only assisting in one aspect of the task assessment process (e.g. recording performance or marking) (Bull & Sharp, 2000a). The first area of the task assessment process that took advantage of computer-support was objective type assessments that automated the marking process (eliminating the marker) and allowed the results to be instantly available. Bull and Sharp (2000b) found that the use of computers to support assessment has many advantages for the assessment process, assessors and students.

Much of the published research in the field of computer-supported assessment relates to higher education, particularly in university settings (e.g. Brewer, 2004), with little specific to school-based education. However, in the school sector assessment of student creative work in the arts has been addressed for some time with, for example, Madeja (2004) arguing the case for alternatives to paper-and-pencil testing for the arts. Further, there has been some research into the use of portfolios for assessment but most often this is for physical, not digital, portfolios. There has been a limited amount of research in the area in Australia, typically these have been small-scale trials in the use of IT to support assessment processes (e.g. Newhouse, 2005). There have also been reports on the use of online testing in Australia, such as by MacCann (2006), but these usually do not involve assessing practical performance and merely replicate paper-and- pen tests in an online environment.

While there has been only limited empirical research into many areas of computer-supported assessment there are many useful theoretical discussions of the issues such as Spector’s (2006) outline of a method for assessing learning in “complex and ill-structured task domains”. While providing useful ideas and rationales these ideas remain largely untested in the reality of classrooms. What is known is that any use of ICT involves school change (Lim & Hung, 2003; Newhouse, Clarkson, & Trinidad, 2005) and will require training of teachers, changes in thinking, and pedagogical understandings that are difficult to take on, even for younger teachers (Newhouse, Williams, & Pearson, 2006).

There has been increasing interest internationally in the application of computer support to improve assessment as indicated in the focus of a recent keynote address by McGaw (2006). The University of Cambridge Local Examinations Syndicate is conducting over 20 projects to explore the impact of new technologies on assessment including using online simulations in assessing secondary science investigation skills (Harding, 2006). Other organisations (e.g. Becta, 2006) or groups of researchers (e.g. Ridgway, McCusker, & Pead, 2006) have reported on exploratory projects, particularly the increasing use of online testing, although rarely for high-stakes assessment and not without some difficulty (Horkay, Bennett, Allen, Kaplan, & Yan, 2006).

The British Psychological Society has produced a set of guidelines for Computer-Based Assessment. While they mainly focus on online testing they provide a conceptual model that includes Assessment Generation, Assessment Delivery, Assessment Scoring and Interpretation, and Storage, Retrieval and Transmission. The latter two were relevant to the present study with the guidelines for developers and users.
2.2 Assessment of Practical Performance

Research in, and the call to investigate “performance-and-product assessment” is not new as pointed out by Messick (1994, p. 14), tracing back at least to the 1960s. However, Messick claims that mainstream schooling showed little interest in this in until an “upsurge of renewed interest” in the 1990s with “positive consequences for teaching and learning” (p. 13). 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. He argues they are social values that require close attention to the intended and unintended consequences of the assessment through considerations of the purposes of the assessment, the nature of the assessed domain, and “construct theories of pertinent skills and knowledge” (p.14). 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.

Lane (2004) claims that in the USA there has been a decline in the use of performance assessments due to increased accountability requirements and resource constraints. She outlines how this has led to a lack of alignment between assessment, curriculum standards, and instructional practices; particularly with regard to eliciting complex cognitive thinking. At the same time Pollitt (2004) argues that current methods of summative assessment that focus on summing scores on “micro-judgements” is “dangerous and that several harmful consequences are likely to follow” (p. 5). Further, he argues that it is unlikely that such a process will accurately measure a student’s “performance or ability” (p. 5), and more holistic judgements of performance are required. Koretz (1998) analysed the outcomes of four large-scale portfolio assessment systems in the USA school systems and concluded that overall the programmes varied in reliability and were resource intensive with “problematic” (p.309) manageability. This body of literature clearly presents the assessment of student performance as critically important but fundamentally difficult with many unanswered questions requiring research.

Globally interest in performance assessment has increased with the increasing use of standards-referenced curricula. Standards-referenced curricula have evolved over the past 20 years particularly from the UK and more recently in Australian states since the early 1990s. The key concept in these curricula was that student achievement was defined in terms of statements describing what students understood, believed or could do. The term standards-referenced has tended to be used recently to indicate that student achievement is measured against defined standards.

Koretz (1998), who defines portfolio assessment as the evaluation of performance by means of a cumulative collection of student work, has figured prominently in USA debate about education reform. He analysed the outcomes of four large-scale portfolio assessment systems in the 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. Koretz concluded that overall the programmes were resource intensive and did not produce “evidence that the resulting scores provide a valid basis for the specific inferences users base on them…” (p.332). 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.

2.3 Digital Forms of Assessment

For the purposes of the study three particular forms of assessment were defined that employed digital technologies to represent the output of student performance. These forms were an Extended Production Exam, a Focussed Performance Tasks Exam and a Reflective Digital Portfolio and were not intended to provide an exhaustive list but rather define major forms that appeared to be relevant to the two courses involved in the study.
2.3.1 Extended Production Exam

An Extended Production Exam was considered to be the completion, under ‘exam conditions’, of one practical assessment task that incorporated a full set of processes (e.g. design process, scientific investigation) and centred on one major scenario. Examples were found locally, nationally and internationally of performance on practical tasks being assessed through an extended production, or small project, under exam conditions. However, most did not involve the use of digital technologies. The most comprehensive example was that of Kimbell et al (2007) in the UK where students spent two consecutive mornings of three hours duration each working on a structured design activity for the production of a pill dispenser. All student work output was collected digitally using a networked Personal Digital Assistant (PDA) device and local server. A relevant national example was the final external assessment for the Design and Technology course in NSW where students submitted a design project. This involved submitting a paper document and a prototype product but was not put in digital form. In Western Australia the final Drama assessment has involved a short individual ‘performance’, that is assessment face-to-face but is also usually videotaped although again this is not typically assessed in a digital form. On a number of occasions over the past decade, samples of Year 3, 5, 7 and 9 students have been assessed in the Monitoring Standards of Education (MSE) programme that has involved completing a short (2 hours in two parts) Technology and Enterprise design brief including prototype production. Once again, typically the final output has been on paper and not in digital format. Also in Western Australia one of the teacher professional associations (Business Educators of WA) has run a competition requiring students to address a design brief in 90 minutes for the Business Information Technology upper secondary course.

2.3.2 Focussed Performance Tasks Exam

A Focussed Performance Tasks Exam was considered to be the completion, under ‘exam conditions’, of a range of practical tasks that are not necessarily logically connected and typically focus on the demonstration of practical skills. Examples were found locally, nationally and internationally of performance on practical tasks being assessed through the completion of short tasks under exam conditions. However, many did not involve the use of digital technologies. In the UK and USA organisations such as the ETS and EdExcel provide online exams of ICT skills. In Victoria students completing a Multimedia VET (Vocational Education and Training) course sit a practical exam using computers. At Edith Cowan University (ECU), for over a decade, in some teacher education units students have completed short practical exams of ICT skills.

In many contexts the use of computer-adaptive tests have been found to be cost-effective, being shorter, more precise and more reliable than other options (Garmire & Pearson, 2006). When these are enhanced with simulations and games many can see the potential to assess areas such as technological competence (Garmire & Pearson, 2006).

2.3.3 Reflective Process Digital Portfolio

A Reflective Process Digital Portfolio was considered to be a collection of digital artefacts of work output with some reflective commentary (journaling) by the student, organised according to specified parameters such as form, structure, and range of samples required. There are many types of digital portfolios used internationally. For this study the portfolios were repositories of previous work-output annotated by the student to explain the inclusion of the artefact and describe its characteristics relevant to assessment criteria. In a review of e-assessment the digital portfolio is recommended as a “way forward” in the high-stakes assessment of “practical” work in that ICT “provides an opportunity to introduce manageable, high quality coursework as part of the summative assessment process (Ridgway, McCusker, & Pead, 2006). Three uses of portfolios are suggested, one of which is “to provide a stimulus for reflective activity”.

Examples were found locally, nationally and internationally of performance on practical tasks being assessed through the submission of a portfolio. However, most did not involve the use of digital technologies. For example, in Western Australia students in the Art course in upper secondary school have traditionally submitted a physical portfolio of work for external assessment. This has also occurred in courses such as Furniture Woodwork and Interactive Media but only at the school level. In the UK Edexcel offers a Diploma in Digital Applications that involves students creating a digital
portfolio (called an eportfolio) of practical work done on computer with teachers required to submit a sample of student portfolios as selected by the authority. While the portfolios did not have a required structure there was a restricted list of file formats. The International Baccalaureate Program also has adopted the use of portfolios for assessment in some areas of study.

Thus, the use of portfolios is not new, particularly in areas such as the visual arts and design and technology but typically these have been paper-based (Garmire & Pearson, 2006). The exercise of assembling a portfolio is often seen as much as a “learning tool” as an “assessment tool” but the results are typically limited by physical storage space and methods of access (Garmire & Pearson, 2006).

**2.4 Methods of Marking**

Task assessment is what is commonly referred to as ‘marking’. Once students have completed the assessment task the output needs to be judged by some method to determine a score, grade or ranking. Three methods of marking are considered here: ‘traditional’ true score marking, judgements using standards-based frameworks, and comparative pairs judgements.

Traditionally summative assessment has tended to involve students ‘sitting’ paper-based exams that are scored by allocating a number to items in the exam and then summing these numbers. This is sometimes called true score marking or cumulative marking. Pollitt (2004) argues that current methods of summative assessment that focus on summing scores on “micro-judgements” is “dangerous and that several harmful consequences are likely to follow” (p. 5). Further, he argues that it is unlikely that such a process will accurately measure a student’s “performance or ability” (p. 5). He claims that this has been tolerated because assessment validity has been overshadowed by reliability due to the difficulty and expense in addressing the former compared with the latter.

Standards-based frameworks and rubrics have been used for many years by teachers in Western Australia and other localities to mark student work but have less often been used for summative high-stakes marking. This involves the definition of standards of achievement against which to compare the work of students. Typically this is operationalised for a particular assessment task through a rubric that describes these standards according to components of the task. The results may be represented as a set of levels of achievement or may be combined by converting these to numbers and adding them. However, using Rasch Modelling they may be combined to create an interval scale score.

Comparative pairs marking involves a number of assessors making judgements on achievement through comparing each student’s work with that of other students, considering a pair of students at a time and indicating the better of the two. This is sometimes referred to as pairwise comparisons or cumulative comparisons.

**2.4.1 Standards Based Frameworks**

In a recent report for the Curriculum Council of WA, Prof Jim Tognolini, states that "One of the main advantages of a standards-referenced assessment system is that the results can indicate what it is students have achieved during the course” and that "at the same time, use the same scores for university entrance purposes”. Further he explains that this provides students with "a meaningful record of their achievements“ and this will "facilitate smoother entry through different pathways into higher education and the workforce". He points out that all Australian states and many international systems including the Baccalaureate and PISA have a standards-referenced curriculum. He defines it as "where educational outcomes are clearly and unambiguously specified" and claims this has "significant power and appeal in more globalised contexts" providing a "mechanism for tracking and comparing outcomes over time and across jurisdictions".

The word "rubric" is a derivative of the Latin word ruber meaning “red”. In literary history, rubrics are margin notes in texts giving description, or common examples for, or about, the passage (Wiggins, 1998). The current research literature on marking keys promotes the use of criterion or rubric based marking keys to enhance transparency, reliability and, when the task is aligned with the learning outcomes, also validity (Andrade, 2005; Tierney & Marielle, 2004). In current usage, a rubric is a guide listing criteria used for rating performance (Wiggins, 1998).
Marking using a rubric based on a standards framework requires assessors to compare a student’s work against what is theoretically possible. This is difficult and requires considerable depth of knowledge and experience and can still result in different assessors judging the same work differently because they have different standards in mind. This leads to a problem of reliability that is typically overcome by using more than one assessor for each piece of work and then having a consensus process. This may be costly and still somewhat unreliable.

Assessment based on a standards framework is not new and has been used in a number of countries for many decades. The best known example is the testing associated with the National Curriculum in the United Kingdom in the 1980s and 1990s. At the school level Schagen and Hutchinson (Schagen & Hutchison, 1994) found that there were a "… variety of different methods used to award Levels based on marks obtained or performance on Statements of Attainment (SoA)." However, at the national level there are a number of National Curriculum Assessment tests that must be completed by all students of selected ages in the UK. Tests of reliability on these tests have found that in some of the National Curriculum Assessment tests "pupils of similar ability could be assigned Levels two or more apart" due to statistical error or other factors such as context, test construction, etc (Schagen & Hutchison, 1994).

In a study titled “Assessing Expressive Learning” that involved nearly 2000 art portfolios and the use of rubrics it was found that “qualitative instructional outcomes can be assessed quantitatively, yielding score values that can be manipulated statistically, and that produce measures that are both valid and reliable estimates of student art performance” (Madeja, 2004).

2.4.2 Comparative Pairs Method of Marking

The comparative pairs judgement method of marking involves Rasch Modelling and was used by Kimbell, Wheeler, Miller and Pollitt (2007) in the E-scape Project delivering an assessor reliability coefficient of 0.93 which is considerably higher than typical coefficients resulting from other methods of marking.

Pollitt (2004) describes the comparative pairs method of marking applied to performance assessment in his paper, “Let’s stop marking exams”. He claims the method he and his colleagues have developed is “in intrinsically more valid” and is “rooted in the psychophysics of the 1920s” (p. 2). He goes on to explain that while the system is better than the traditional system to this stage, it has not been feasible to apply, due to time and cost constraints, however, with the use of ICT to support this system these constraints are removed and “Thurstone’s methods” that “have waited 80 years … are at last … feasible” (p. 21). He quotes Laming that there is “no absolute judgement. All judgements are comparisons of one thing with another” and explains that it is more reliable to compare performances or products between students than with “descriptions of standards” (p.6). He claims that they have more than ten years experience in applying the method in a variety of contexts and that with expert application about 20 comparisons per student is required. However, he does suggest that the method should not be used with every type of assessment, with research required to determine the appropriateness and whether “sufficient precision can be achieved without excessive cost” (p. 16).

McGaw (2006) also believes that the comparative pairs method of marking provides an opportunity to improve the validity of high-stakes assessment in separating the “calibration of a scale and its use in the measurement of individuals” (p. 6). He claims that while the “deficiency” of norm-referenced assessment has been understood for many years it was seen that there was no alternative. Now he believes that with methods involving comparisons being supported of digital technologies there is an alternative that should be explored.

An important question is whether the advances in psychometrics that permit calibration of scales and measurement of individuals that allows interpretation of performance in terms of scales can be applied in public examination. (McGaw, 2006, p. 7)

2.5 Digital Representations of Assessment Tasks

In order to judge student performance that performance needs to either be viewed or represented in some form. This may involve the assessor viewing a student performing, such as in a musical recital, or viewing the results of a student performing, such as in an art exhibition. Most often the latter occurs
because this is either more appropriate or more cost-effective. In places such as Western Australia the inclusion of either type of assessment for high-stakes purposes has been rare due to the costs and logistics involved. For example, student performance in conducting science experiments has not been included because of the difficulty in supervising the students and viewing their work, and production in design and technology, or home economics related areas, has not been included because the products are bulky and therefore difficult to access by assessors. However, many forms of student performance can be recorded in digital representations using video, audio, photographic or scanned documents, and some student work is created in digital format using computer software. In these cases the representations of student work can be made available to assessors relatively easily and cheaply using digital repositories and networked computer workstations.

A ground-breaking project aimed at assessing performance, titled E-scape, is being conducted by the Technology Education Research Unit (TERU) at Goldsmiths College, University of London (Kimbell, Wheeler, Miller, & Pollitt, 2007). This project built upon many years of work on improving assessment in the design and technology curriculum (Kimbell, 2004). E-scape combines three innovations in the assessment of practical performance by representing student work entirely in digital form, collating this work using an online repository, and marking it using a comparative pairs judgements technique. Their results have been encouraging with student work being assessed on an interval scale with a reliability coefficient of 0.93.

Kimbell developed a feasibility framework consisting of four dimensions useful for evaluating digital forms of assessment. This framework is described below.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manageability</td>
<td>Concerning making a digital form of assessment do-able in typical classrooms with the normal range of students.</td>
</tr>
<tr>
<td>Technical</td>
<td>Concerning the extent to which existing technologies can be adapted for assessment purposes within course requirements.</td>
</tr>
<tr>
<td>Functional</td>
<td>Concerning reliability and validity, and the comparability of data with other forms of assessment.</td>
</tr>
<tr>
<td>Pedagogic</td>
<td>Concerning the extent to which the use of a digital assessment forms can support and enrich the learning experience of students.</td>
</tr>
</tbody>
</table>

**2.6 Guidelines for Design and Development of Digital Assessments**

As in most areas of education, and particularly for assessment, authorities and/or researchers in many localities have developed guidelines for the use of digital technologies with assessment processes. For example, the British Psychological Society published a set of general guidelines for the use of “Computer-Based Assessments” through its Psychological Testing Centre (The British Psychological Society, 2002). These guidelines include the use of digital technologies in Assessment Generation, Assessment Delivery, Assessment Scoring and Interpretation, Storage, Retrieval and Transmission. These guidelines are defined from a developer and user perspective. Similarly, The Council of the International Test Commission developed international guidelines for good practice in computer-based and Internet delivered testing (The Council of the International Test Commission, 2005). These were focussed on four issues: the technology, the quality of the testing, the control of the test environment, and the security of the testing. The contexts considered all involved students sitting at a computer to complete a test.

All assessment items are required to be valid, educative, explicit, fair and comprehensive, and should allow for reliable marking. The descriptions of the digital assessment tasks below assume this but focus on any areas that are of a particular challenge to that assessment type.

**2.6.1 Guidelines Specific to Computer-Based Exams**

It was likely that, in some cases in the project, computer-based exams would be used for the Applied Information Technology course. The International Test Commission has provided detailed guidelines for computer-based exams (The Council of the International Test Commission, 2005). These
guidelines were specific to test developers, test publishers and users. An array of specific guidelines was presented according to the following structure.

(1) Give due regard to technological issues in Computer Based Testing (CBT) and Internet testing
   a. Give consideration to hardware and software requirements
   b. Take account of the robustness of the CBT/Internet test
   c. Consider human factor issues in the presentation of material via computer or the Internet
   d. Consider reasonable adjustments to the technical features of the test for candidates with disabilities
   e. Provide help, information, and practice items within the CBT/Internet test

(2) Attend to quality issues in CBT and Internet testing
   a. Ensure knowledge, competence and appropriate use of CBT/Internet testing
   b. Consider the psychometric qualities of the CBT/Internet test
   c. Where the CBT/Internet test has been developed from a paper and pencil version, ensure that there is evidence of equivalence
   d. Score and analyse CBT/Internet testing results accurately
   e. Interpret results appropriately and provide appropriate feedback
   f. Consider equality of access for all groups

(3) Provide appropriate levels of control over CBT and Internet testing
   a. Detail the level of control over the test conditions
   b. Detail the appropriate control over the supervision of the testing
   c. Give due consideration to controlling prior practice and item exposure
   d. Give consideration to control over test-taker’s authenticity and cheating

(4) Make appropriate provision for security and safeguarding privacy in CBT and Internet testing
   a. Take account of the security of test materials
   b. Consider the security of test-taker’s data transferred over the Internet
   c. Maintain the confidentiality of test-taker results

Clearly many of the guidelines apply generally to any test-taking context (e.g. 2d, 2e and 2f), whether on computer, or not. Many of the other guidelines were not applicable to the current project (e.g. 4a, b and c) because only single classes and their teachers in particular schools were involved. However, many of the guidelines in the first three areas were relevant to one or more of the cases in the project. For example, in all cases involving computer-based exams some of the guidelines associated with 1a, 1b, 2a and 2b were relevant, and to some extent some guidelines associated with 3a, 3b and 3d were relevant. Even so they were mainly relevant to the implementation of large scale online testing.

**2.6.2 Guidelines Specific to Production Exams**

Production exams would not necessarily be computer-based, for example, production exams in design and technology need only be represented digitally through records of the performance (e.g. video, photograph, scanned document). The areas of concern with production exams are; ensuring that they are fair to all students in terms of access to information, materials, and tools, that they are valid in assessing what is intended, and provide for reliable marking given the usually varied types of student work output. Therefore it is often recommended that the assessment task be well bounded, the work environment be limited (e.g. access to a limited set of information or tools), the time available be controlled, student work be invigilated, and the required work output be well defined.
2.6.3 Guidelines Specific to Performance Tasks Exams
Performance tasks exams would not necessarily be computer-based with, for example, exams in design and technology only being represented digitally through records of the performance (e.g. video, photograph, scanned document). For performance task exams it is recommended that the tasks be clearly defined and limited, the work environment be narrowly prescribed (e.g. access to prescribed information or tools), and the required work output be well defined.

2.6.4 Guidelines Specific to Digital Portfolios
The main concerns with the use of digital portfolios for assessment are:

• The authentication of student work given the period of time within which work is completed
• Ensuring that they are fair to all students in terms of access to information, materials and tools
• That they can be marked reliably given the usually varied types of student work output.

Therefore it is often recommended that the portfolio require a particular structure and limit the contents in type and size, the time available be controlled, and the work be authenticated by a teacher and the students. In a review of e-assessment it was suggested that a digital portfolio may involve three sections: student self-awareness; student interaction; and thinking about futures and informed decisions (Ridgway, McCusker, & Pead, 2006). In British Columbia, Canada, students complete a graduation portfolio. They are provided with a number of guides as Word documents that act as templates to construct their portfolios.

Carney (2004) developed a set of critical dimensions of variation for digital portfolios:
1. Purpose(s) of the portfolio;
2. Control (who determines what goes into the portfolio and the degree to which this is specified);
3. Mode of presentation (portfolio organisation and format; the technology chosen for authoring);
4. Social Interaction (the nature and quality of the social interaction throughout the portfolio process);
5. Involvement (Zeichner & Wray identify degree of involvement by the cooperative teacher important for preservice portfolios; when considered more broadly, other important portfolio participants might include university teachers, p-12 students and parents, and others); and
6. Use (can range from low-stakes celebration to high-stakes assessment).

Barrett (2005) suggests the following.

Identify tasks or situations that allow one to assess students’ knowledge and skills through both products and performance. Create rubrics that clearly differentiate levels of proficiency.
Create a record keeping system to keep track of the rubric/evaluation data based on multiple measures/methods. Provide opportunities for students to learn and resubmit, maximizing diagnosis and remediation. Model the power of assessment as learning. (p. 10)

She goes on to suggest that for “Portfolios used for Assessment of Learning” that is for summative assessment the following are defining characteristics.

• Purpose of portfolio prescribed by institution
• Artefacts mandated by institution to determine outcomes of instruction
• Portfolio usually developed at the end of a class, term or program - time limited
• Portfolio and/or artefacts usually "scored" based on a rubric and quantitative data is collected for external audiences
• Portfolio is usually structured around a set of outcomes, goals or standards
• Requires extrinsic motivation
• Audience: external - little choice
2.7 Conceptual Framework for the Study

In order to investigate the use of digital representations to deliver authentic and reliable assessments of performance this study brought together three key innovations:

1. The representation in digital files of the performance of students doing practical work.
2. The presentation of digital representations of student performance in an online repository so that they are easily accessible to markers.
3. Assessing the digital representations of student performance using both standards-based judgement and the paired comparison judgement methods with holistic and component criteria-based judgements.

While each of these innovations is not new in themselves, their combination applied at the secondary level of education is new. Apart from Kimbell’s (2007) work at the University of London there was no known precedent.

Fundamentally this study investigated the use of digital forms of representation of student practical performance for summative assessment, whether the student created digital files or their performance was recorded in digital format by filming, photographing, audio recording or scanning. The overall study investigated three digital forms of assessment: Extended Production Exam, Focussed Performance Tasks, and Reflective Portfolio.

**Extended Production Exam:** the completion, under exam conditions, of one common assessment task that incorporates a full set of processes (e.g. design process, scientific investigation) and centres on one major scenario for which students are required to submit all output in digital formats.

**Focussed Performance Tasks:** a range of focussed practical tasks that students submit in digital formats after working under exam conditions. This may be at a computer or may be a digital recording of their performances.

**Reflective Process Portfolio:** a collection of digital artefacts of work output completed by the student, according to specified parameters such as form, structure, and range of samples required. These will require some reflective commentary by students.

The digital representations of student performance were combined within an online repository. The use of online repositories for student work output is increasingly common, often referred to as online portfolios, with many products available to facilitate their creation and access (Richardson & Ward, 2005). The key feature is that the portfolios can be accessed from anywhere and thus markers from different jurisdictions can be involved, enhancing consistency of standards.

The paired comparison judgement method of marking, involving Rasch Modelling, was trialed using holistic and component criteria judgements for one case study school. While Pollitt (2004) describes the method as “inextricably more valid” and better than the traditional system, he believes that without some ICT support it has not been feasible to apply due to time and cost constraints, and he does suggest that further research is required to determine the appropriateness and whether “sufficient precision can be achieved without excessive cost” (p. 16). McGaw (2006) believes that such methods being supported by digital technologies should be applied in public examinations.

The diagram in Figure 2.1 represents the main concepts involved in assessment with the study focussing initially on the Assessment Task and thereby the Method of Assessment and the Student Work itself. These components are highlighted in Figure 2.1. However, to investigate the achievement of the desired performance indicators that relate to Kimbell’s feasibility framework the study was also involved with Task Assessment, in particular marking activities using standards frameworks and the use of the comparative pairs marking approach.
Figure 2.1: A diagrammatic representation of a conceptual framework for the assessment of performance. (Based on the work of Campbell (2008).)
3 - Method

The general aim of this study was to explore the potential of various digitally based forms for external assessment for senior secondary courses in Western Australia. Specifically there was a need to determine the cost effectiveness of each format in terms of the need to support an outcomes-based curriculum framework for students in schools across the state. The problem being addressed was the need to provide students with assessment opportunities in new courses 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. That is, the external assessment for a course needs to accurately and reliably assess the outcomes without a huge increase in the cost of assessment.

Therefore the specific aims of the research were:

1. To identify the benefits of each digitally based form to support the external assessment of outcomes in a senior secondary course.
2. To identify constraints to the implementation of each digitally based form to support the external assessment of outcomes in senior secondary courses in typical school settings.
3. To compare the benefits and constraints between each digitally based form to support the external assessment of outcomes in senior secondary courses in order to determine the suitability of each for the different types of courses.

The study was a ‘proof of concept’ project to explore the feasibility of a range of digitally-based formats for external assessment. The feasibility was investigated within a framework consisting of the four dimensions: technological, pedagogic, manageability, and functionality. These dimensions were investigated through trials of the formats in sample schools with sample students in two courses: Applied Information Technology, and Engineering. This built on the work by Kimbell and Wheeler (2005) from the United Kingdom.

The study was evaluative in nature set within an ethnographic framework in that activity was considered to occur within learning environments where the characteristics of teachers and students and the culture created are critical to an understanding of all aspects of the curriculum and pedagogy, including assessment. Therefore, this project employed an ethnographic action research evaluation methodology using interpretive techniques involving the collection of both qualitative and quantitative data. The study drew on the traditions of interpretive research but also sought to employ, where appropriate, the quantitative methods of more traditional positivist research. That is, quantitative measures concerning achievement, costs and resource use were used but need to be interpreted within the context of the learning environment and in which the measures occurred.

The main research question became:

How are digitally based representations of student work output on tasks most effectively used to support highly reliable summative assessments of student performances for the Applied Information Technology and Engineering senior secondary courses?

The research design can be described as participative action research with participants contributing to development through evaluative cycles. As such this required an analysis of the perspectives of the key groups of participants (teachers, assessors, students) with data collected from each group. These data were compiled into case studies within a multi-case approach (Burns, 1996) in which each case is defined by one digital form of assessment in one class for one course. This approach allowed for refinement and further development of findings based on multiple instances of the same phenomenon under different conditions (Willig, 2001). Therefore, this study largely employed an ethnographic action research evaluation methodology using interpretive techniques involving the collection of both qualitative and quantitative data.

The study focused on the Applied Information Technology (AIT) and Engineering Studies courses and sought to involve as many schools, teachers and students as possible. The AIT course was selected because at least two of its outcomes were directly related to the production of digital materials and students and teachers were more likely to have adequate competence in the use of ICT to be able to readily implement a range of digitally-based formats of assessment. The Engineering course was
selected because it is a completely new course and its outcomes include processes and practical performance. However, application to many other courses was considered when drawing up the conclusions.

3.1 Samples
The study involved 14 teachers and their upper secondary students in either Applied Information Technology (AIT) or Engineering. There were seven teachers for AIT and seven for Engineering.

<table>
<thead>
<tr>
<th>Group</th>
<th>Course</th>
<th># Tchrs</th>
<th># Studs</th>
<th>Digital Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>AIT: Year 11 (1A/1B)</td>
<td>1</td>
<td>15</td>
<td>Portfolio - slideshow</td>
</tr>
<tr>
<td>CA</td>
<td>AIT: Year 11 (2A/2B)</td>
<td>1</td>
<td>10</td>
<td>Exam (Production) – web page</td>
</tr>
<tr>
<td>IA</td>
<td>AIT: Year 12 Interactive Media</td>
<td>1</td>
<td>15</td>
<td>Portfolio – Flash movie</td>
</tr>
<tr>
<td>MA</td>
<td>AIT: Year 12 Industry IT</td>
<td>1</td>
<td>9</td>
<td>Exam (Production) - video</td>
</tr>
<tr>
<td>PA</td>
<td>AIT: Year 11 (2A/2B)</td>
<td>2</td>
<td>17</td>
<td>Portfolio – various formats</td>
</tr>
<tr>
<td>WA</td>
<td>AIT: Year 11 (2A/2B)</td>
<td>1</td>
<td>10</td>
<td>Portfolio – web page</td>
</tr>
<tr>
<td>XA</td>
<td>AIT: Year 11 (2A/2B)</td>
<td>1</td>
<td>48</td>
<td>Exam (Tasks) - poster</td>
</tr>
<tr>
<td>GE</td>
<td>Year 11 – Eng 2A/B</td>
<td>1</td>
<td>13</td>
<td>Portfolio: vehicle design</td>
</tr>
<tr>
<td>ME</td>
<td>Year 11 – Eng 2A/B</td>
<td>1</td>
<td>9</td>
<td>Production exam</td>
</tr>
<tr>
<td>VE</td>
<td>Year 11 – Eng 2A/B</td>
<td>1</td>
<td>9</td>
<td>Performance tasks exam</td>
</tr>
<tr>
<td>WE</td>
<td>Year 12 – Eng 3A/B</td>
<td>1</td>
<td>10</td>
<td>Production exam</td>
</tr>
<tr>
<td>HE</td>
<td>Year 11 – Eng 2A/B</td>
<td>1</td>
<td>19</td>
<td>Portfolio</td>
</tr>
<tr>
<td>YE</td>
<td>Year 11 – Eng 2A/B</td>
<td>1</td>
<td>15</td>
<td>Performance tasks exam</td>
</tr>
</tbody>
</table>

A multi-case approach involving multiple instances of classes of students completing digitally based tasks to be assessed using an outcomes based framework was used. This approach allows for refinement and further development of findings based on multiple instances of the same phenomenon under different conditions.

The study involved a sample of classes of students currently enrolled in at least one of the subjects Interactive Media, Industry Information Technology and Business Information Technology from at least ten schools. Data from all sources on each class was combined to create a series of case studies structured to highlight design and implementation features.

An Expert Reference Group of representatives from the school systems and experts in assessing outcomes in technology education was assembled to work with the researchers to ensure a strong connection between schools, assessment theory and the study.

3.2 Data Collection and Analysis
A range of types of quantitative and qualitative data were collected including observation in class, a survey of students, a survey of the teacher, interviews with the teacher and a group of students, student work output from the assessment task, and the assessment records of the teacher.

3.2.1 Digital Work Output
Output from student work on the assessment tasks was collected in digital form and placed in the online digital repository to be available to the assessors. The assessment task could not be standardised and thus was quite different for each school/class. Thus there were differences in parameters such as time constraints, level of supervision, level of scripting for implementation, minimum requirements for equipment/resource availability, and types of data to collect. The students’ work was assessed by at least two external assessors and for one case a panel of five assessors, in addition to the classroom
teacher, using a set of criteria developed for each assessment task. For each case a mark and ranking were created for each assessor and a consensus mark and ranking. These were compared using a test of correlation. For the case involving a panel of five assessors a comparative pairs method of marking was used to generate an overall mark and ranking for each student. This was done in addition to the standard two external assessor strategy.

3.2.2 Achievement Data
Teachers provided achievement data for their students for the assessment task and for other school-based assessments. These provided a measure of convergent validity. These data were provided as scores, percentages, ranks or grades. For each case these data were used to create a ‘teacher mark’ and ‘teacher ranking’ for the students. Where grades were provided these were converted to marks using conversion tables developed by the researchers for that case. The teacher mark and teacher ranking were compared with the external assessors’ marks and rankings using a correlation test.

3.2.3 Interviews
Interviews were used to elicit the experiences of students in completing assessment tasks, teachers in supporting them in these, and assessors in marking the student work to address the manageability, functional and pedagogic dimensions of feasibility. A stratified sample of students in each class was interviewed as a group, and all teachers, as soon after completion of the assessment tasks as possible. The student interviews were conducted by one member of the research team and were digitally recorded and recorded in note form on structured interview schedules. The combination of notes and digital audio provided a complete record of the interview. The teacher interviews were either conducted face-to-face or using email by one of the researchers, based on structured interview schedules. Assessor interviews were conducted in an adhoc manner by one of the researchers.

The student interviews were transcribed and then summarised in the case study reports so that themes could be identified. The data from the teacher and assessor interviews were summarised and included in the case study reports (see sections 4 and 5 of this report).

3.2.4 Observations
Observations of each class of students in the process of completing the assessment task occurred using a structured approach to address the manageability dimension of feasibility. These data assisted in interpreting results from other data particularly in terms of the constraints associated with the realities of conducting these assessments in schools. Notes were written or recorded during observation periods and were verified by the participating teacher as soon as possible after the observation. A summary was included in the case study reports.

3.2.5 Student Questionnaire
Questionnaires were employed to collect data from students on their characteristics, including perceived level of ICT skills and experience, and their experience of the assessment task. These data were used to address the manageability dimension of feasibility. The questionnaires were administered to all students from the sample classes as soon after the assessment tasks as possible. The data were entered into a spreadsheet and SPSS. Descriptive and frequency statistics were generated to include in the case study reports.

3.2.6 Development of Case Studies
These quantitative and qualitative data for each class were compiled into case studies with a report generated for each case study. Each case study report was sent to the teacher(s) involved for validation and some interpretation. Triangulation of data types and sources enhanced the credibility of findings. Validity and authenticity of the assessment tasks was judged in relation to three criteria: how well the performance of students matches the curriculum outcomes; the extent to which the method of representing performance is authentic; and, whether the task and context are meaningful and relevant to students and community practice. The outcomes from each case study were then summarised according to the dimensions of feasibility so that the data could be analysed using a constant comparative approach looking for themes, trends and developing rich descriptive accounts (Patton,
Data were coded according to emergent themes. Themes were constantly compared with emergent categories to establish a best fit with the data set.

3.3 Methodology Framework

The study included three main phases: Exploration of the Alternatives, Development of the Prototypes, Implementation and Evaluation of the Prototypes. In Table 1.1 these phases are expanded to give the methods of data collection. All data were analysed by looking for themes, trends and for descriptive purposes. Resulting data sets were analysed using a constant comparative approach (Patton, 1990). This collaborative research project was designed to be the first phase of a larger research and development study conducted over a period of at least three years with three main phases: Investigation and Development, Implementation, and Summative Evaluation. The present study was the Investigation and Development phase that took one year with the major outcomes being a comprehensive literature review and the development and testing of a number of potential prototypes for the AIT and Engineering courses.

Table 1.1
Data sources and planned methods for each phase of the study.

<table>
<thead>
<tr>
<th>Purpose / Description</th>
<th>Data source</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exploration of the Alternatives</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literature Review</td>
<td>Journals, reports and anecdotes</td>
<td>A paper was developed (the previous chapter) outlining guidelines for design and development of assessment forms.</td>
</tr>
<tr>
<td>Evaluation of alternatives</td>
<td>Reference Group of Experts</td>
<td>Separate design briefs for assessment prototypes were developed for AIT and Engineering.</td>
</tr>
<tr>
<td>Semi-structured Interviews</td>
<td>Selected teachers</td>
<td>The data was analysed by describing the communication with the teacher for each case.</td>
</tr>
<tr>
<td><strong>Development of the Prototypes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customised Prototypes</td>
<td>Research team</td>
<td>The degree of customisation of features and functionality of assessment forms in the development of the prototype is described for each case.</td>
</tr>
<tr>
<td>Evaluation of Designs</td>
<td>Research team</td>
<td>For each case there was an examination of the prototype to judge compliance with recognised usability principles.</td>
</tr>
<tr>
<td><strong>Implementation and Evaluation of the Prototypes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom Observation</td>
<td>Researcher notes from visits</td>
<td>Determine constraints to implementation of each prototype within the four dimensions of feasibility.</td>
</tr>
<tr>
<td>Student Questionnaire</td>
<td>Responses to questionnaire</td>
<td>Quantitative statistical analysis to test for factors relating to relative success of the assessment task for each student.</td>
</tr>
<tr>
<td>Achievement of students</td>
<td>Teachers and Expert Assessors Research team</td>
<td>An evaluation of validity and reliability of the results from the prototypes.</td>
</tr>
<tr>
<td>Semi-structured interviews</td>
<td>Selected teachers</td>
<td>Information from the teachers within the four dimensions of feasibility about: usability, difficulties, attitudes, feelings and reactions.</td>
</tr>
</tbody>
</table>
3.3.1 Research Tasks
Teachers were given the following instructions as to the research tasks that would be undertaken. These were chronological and divided into two sets.

Prior to and During the Assessment Task
(1) A member of the research team will work with you in planning the assessment task.
(2) Collect parent permission forms for all students.
(3) The member of the research team will visit the school to observe the students working on the assessment task that you will administer.

After the Assessment Task
(4) You will assess the student work for the purposes of their course in whatever manner you see fit. You will need to provide us with a class list of results (names removed).
(5) A member of the research team will collect copies of the students' work off you.
(6) A member of the research team will administer a questionnaire with your students.
(7) A member of the research team will interview a group of 4 or 5 students in a forum.
(8) You will receive the results of marking by the independent assessors.
(9) You will receive a short questionnaire to return via email with some follow-up phone or face-to-face questions.
(10) You will be asked for feedback on a short report on your class's work.

3.3.2 Working with Teachers
The research team worked with each teacher on one of their assessments to do the minimum modification to match it with one of the three assessment formats. The researchers did the additional work ensuring the approval of the teacher concerned before finalising the assessment. The teacher administered the assessment with a researcher observing (where necessary technical support was provided by the research team).

The teacher assessed the student work for the purposes of their course and then at least two researchers assessed the work anonymously for the purposes of the research. At all times the aim was to follow the principle of minimum intrusion.
### 3.3.3 Final Timeline of Procedures for the Study

Table 1.2 presents a timeline of the procedures involved in the study.

**Table 1.2**

Timeline of Procedures for the Study

<table>
<thead>
<tr>
<th>Month(s)</th>
<th>Procedure</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 2006</td>
<td>Seek ethics clearance</td>
<td>Team leader</td>
</tr>
<tr>
<td>August</td>
<td>Complete review of literature originally commenced in 2005.</td>
<td>Team</td>
</tr>
<tr>
<td>September</td>
<td>Select samples – teachers, classes in Year 11/12 studying Applied Information Technology and subjects similar to Engineering. Interview participating teachers to identify important features of their programmes and the ICT infrastructure of the school.</td>
<td>Team</td>
</tr>
<tr>
<td>October-March 2007</td>
<td>Design and creation of assessment customized prototypes. Evaluate designs Intend to include Prof Richard Kimbell (UK) as a visiting scholar.</td>
<td>Team Teachers</td>
</tr>
<tr>
<td>April-July</td>
<td>Trial of prototype assessments with sample classes. Survey of students involved. Classroom observations, interviews with teachers and students.</td>
<td>Teachers Team</td>
</tr>
<tr>
<td>August-December</td>
<td>Data analysis – achievement, survey &amp; interview, observation</td>
<td>Team</td>
</tr>
<tr>
<td>December-January 2008</td>
<td>Review and interpretation of results</td>
<td>Teachers Team</td>
</tr>
<tr>
<td>February-March</td>
<td>Compile findings and identify topics for publications</td>
<td>Team</td>
</tr>
<tr>
<td>April</td>
<td>Complete and disseminate project report.</td>
<td>Team</td>
</tr>
</tbody>
</table>
4 - Applied Information Technology Case Studies

There were seven schools, eight teachers and nine classes of students involved in the project associated with the Applied Information Technology course, however, two of the teachers were delivering the older subjects to Year 12 students while the others were delivering the new course to Year 11 students. Seven case studies were constructed and are reported upon in this section. The case studies are briefly introduced below.

<table>
<thead>
<tr>
<th>Case</th>
<th>Type of Assessment</th>
<th>Classes and Courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>WA</td>
<td>Digital portfolio: web-based template</td>
<td>Year 11 – AIT 2A/B</td>
</tr>
<tr>
<td>AA</td>
<td>Digital portfolio: slideshow</td>
<td>Year 11 - AIT 1A</td>
</tr>
<tr>
<td>IA</td>
<td>Digital portfolio: Flash movie</td>
<td>Year 12 – Interactive Media</td>
</tr>
<tr>
<td>PA</td>
<td>Digital portfolio: Choice of format</td>
<td>Year 11 – AIT 2A/B</td>
</tr>
<tr>
<td>MA</td>
<td>Digital video production exam (2 weeks)</td>
<td>Year 12 – Industry IT</td>
</tr>
<tr>
<td>CA</td>
<td>Web page production exam (3 hours)</td>
<td>Year 11 – AIT 2A/B</td>
</tr>
<tr>
<td>XA</td>
<td>Poster graphics task exam (50 mins)</td>
<td>Year 11 – AIT 2A/B</td>
</tr>
</tbody>
</table>
WA: Website portfolio based on template
The WA case study involved one class of Year 11 students at one school completing the Applied Information Technology course Unit 2A. There were 15 students in the class that was in a Business/Multimedia context, but five students were not involved in the study, because either they did not want to be involved or their parents did not give permission.

The aim was for each student to collate a portfolio of work stored in digital formats, emanating from work completed in their course, and collated as a website using a template provided. The template provided a structure related the four outcomes of the course and the tasks completed. There were FIVE tasks (Task 4A and 4B were treated as different tasks) completed throughout the semester and students were required to include annotations for each of the outcomes. The portfolio addressed Outcomes 1, 2, 3 and 4 of course and was assessed using the related levels of achievement. The teacher developed the tasks and portfolio template based on materials provided by the Curriculum Council. The researcher visited the classes to observe them working on their portfolio.

Students were required to select digital artefacts and process documents from their FIVE TASKS and some classroom exercises that demonstrated their highest achievement for each of the FOUR OUTCOMES. The emphasis was on quality not quantity although students tended to include everything they had done at that stage. It contributed to the assessment of the student’s ability to use and evaluate the application of ICT relevant to personal, community, business and industry settings. The portfolio provided evidence for experience in which they have developed an understanding of, and practical skills in the use of ICT relevant to these settings. Each digital file should have included an annotation indicating how it was created, the format of the file, and an explanation of the purpose and uses of the file. However, this was rarely done. The teacher facilitated the collation of the portfolio.

Results from data analysis
A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class
Towards the end of the first semester the 10 students were observed working on the portfolio in a computer laboratory for one hour. They used a web template that provided a structure to import work and include annotations around the Four Outcomes and Five Tasks. All annotations were within a single page with links to original material from the tasks. Most students were just beginning to add links and annotate, some had not finished the tasks. The teacher was concerned about Outcome 2 and whether the students had anything relevant from the tasks. It was suggested that the ‘Peer Evaluation’ in Task 2 may be suitable. The students were using Macromedia Dreamweaver to edit their website.

Survey of Students
There were 8 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.
Table 4.1

Descriptions and descriptive statistics for the scales based on items from the student questionnaire (N=8).

<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>8</td>
<td>2.9</td>
<td>3.6</td>
<td>3.1</td>
<td>0.22</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>8</td>
<td>1.4</td>
<td>2.6</td>
<td>2.1</td>
<td>0.37</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>8</td>
<td>2.0</td>
<td>2.6</td>
<td>2.3</td>
<td>0.26</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confidence</td>
<td>8</td>
<td>2.0</td>
<td>2.7</td>
<td>2.4</td>
<td>0.24</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>8</td>
<td>2.6</td>
<td>3.5</td>
<td>3.2</td>
<td>0.25</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>8</td>
<td>0</td>
<td>106</td>
<td>75.8</td>
<td>37.47</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

Figure 4.1: Distribution of scores for the Skills and SCuse scales on the student questionnaire.

The students didn’t tend to have done the nominated activities in Q10 and were not confident they could do them except by typing an assignment for school that they all had done. Most liked using computers at school and home but only sometimes liked finding out things for themselves. Seven of the eight indicated great confidence in using computers. Six had Broadband Internet access at home and five used computers on a daily basis.

Almost all Agreed or Strongly Agreed with Q2 items concerning value of eAssessment (only two disagreed on any of these items). Most had little experience with creating a digital portfolio but felt that they could quickly get used to it.

Students indicated that the assessment was relatively easy to complete, that they typically had done the range of applications indicated, that they were confident in using ICT, and had a high level of ICT skills across the range of types of applications. On average they indicated using ICT for about one and quarter hours per day at school with the lowest 0 minutes and the highest 106 minutes.

There were two open-ended questions on the student questionnaire. Item 3 asked students to list the two best things about the assessment and Item 4 asked them to list the two worst things about the assessment. Generally students considered that using computers made it easier and faster, was a better environment, allowed access to more information and tools and was easy to manage and move things around. Main worst things were loss of data and, even though experienced with using computers, they had difficulties with using new software.
Students Forum
The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.

The students explained that they created a digital portfolio, using an Internet based program, Macromedia Dreamweaver, to create several pages of outcomes, telling people about what they had done. They put the pieces of work together that they had done over the two terms onto this web page. They had done the creative designs for small business, integrating ICT and the first project was to add ICT into a small businesses, a type of system analysis. They felt that doing the portfolio was “sort of OK”, not terribly hard or easy, and that other students were mostly OK, although some didn’t like doing it.

They didn’t really want to change anything because, it worked out fine. They felt that Dreamweaver was a suitable tool to do the portfolio because the operation was just dragging, clicking and no need to code if you didn’t want to. They believed that they had enough time, with two terms to do it, starting from scratch. This was the first time they had done this type of activity.

They could not identify any technical problems and felt that they had everything they needed. However, there were some access problems with banned websites based on words, one word wrong and it was banned. However, there were enough computers for everyone, although sometimes they were a bit slow when everybody was on the Internet. They would prefer to work on the digital portfolio rather than an exam.

Structured Interview with Teacher
The teacher believed that the task gave the students numerous opportunities to meet the requirements of the outcomes and the opportunity to express themselves and take ownership of the activity. He felt that the activity was well structured and gave the students opportunity to meet the requirements of the outcomes, however given the students prior knowledge was unable to commit enough time. He felt that initially the students found it very difficult as they had a lot of new information to deal with but that their understanding had grown so that they were getting a lot more out of the activity towards the end. They did not experience any technical or other types of problems in implementing the portfolio.

The teacher was not happy with the quality of work the students produced that was below average because at the time of submitting the work the students had not spent enough practical time on the computers to create something of a high standard. While one or two of the students performed better than the rest of the class overall they did not perform to his expectations. He was sure that the students were excited and encouraged by the project and the work they were producing. He believes that portfolio use should be encouraged more in schools and could be used in any subject with the only limitations being access to technology and teachers knowledge of technology.

Results of Marking
The portfolio output for 10 students were collected on a DVD, copied from the school’s server upon which the students saved their work. Each student had a folder containing folders and files that made up their portfolio. The formats of the files included HTML, DOC, PUB, and PPT. Before marking could commence all the portfolios were checked, zipped and uploaded into the MAPS online portfolio system.

The students’ work was marked by two external assessors using a standards-based rubric. At the same time the teacher marked the students’ work using his own analytic marks-based system. The two external assessors marked the student work on the criteria associated with the four outcomes of the course while the teacher awarded a level for each outcome that was converted to a mark out of 100 with 25 allocated for a level 4 for each outcome. The assessor judgements on the outcome was converted to numbers (ranging from 12 to 25) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0.
For the standards-based marking method the assessors used the rubric provided in the course document based on the four outcomes, as described below. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the following instructions on the criteria to assess.

<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Technology Process: Students apply a technology process when creating or modifying products using information communication technologies.</td>
<td>Items from Research Report (Task 1) and Character Design (Task 2). Annotations under Outcome 1 Task 2: characters investigated, criteria developed to evaluate characters, developer &amp; user needs.</td>
</tr>
<tr>
<td>1.1 investigate ideas, and evaluate alternative designs and proposals</td>
<td>Items from Research Report (Task 1) and Character Design (Task 2). Annotations under Outcome 1 Task 2: diagram of character</td>
</tr>
<tr>
<td>1.2 devise communicate and evaluate proposals and design plans in appropriate forms</td>
<td>Items from Research Report (Task 1) and Character Design (Task 2). Annotations under Outcome 1 Task 2: self evaluation, peer evaluation of other, improvements, and timeline.</td>
</tr>
<tr>
<td>1.3 Implement and evaluate production processes and strategies to manage resources efficiently</td>
<td></td>
</tr>
<tr>
<td>2 - Understanding Information and Communication Technologies: Students understand the nature and use of computer hardware and software to achieve information solutions.</td>
<td>Items from Research Report (Task 1) and Healthy Lifestyle Promotion (Task 4B).</td>
</tr>
<tr>
<td>2.1 understand the ICT-related concepts, standards and terminology required to select and use appropriate computer software and hardware to achieve information solutions</td>
<td>Items from Research Report (Task 1) and Healthy Lifestyle Promotion (Task 4B).</td>
</tr>
<tr>
<td>2.2 understand the standards and conventions when selecting and using ICT systems to achieve information solutions</td>
<td>Items from Research Report (Task 1), Character Design (Task 2) and Healthy Lifestyle Promotion (Task 4B). Task 2: Evaluations may refer to production techniques.</td>
</tr>
<tr>
<td>2.3 understand the management, processes, functions, types and relationships between the components of ICT systems required to achieve information solutions.</td>
<td></td>
</tr>
<tr>
<td>3 - Quality of Information Solutions: Students explore alternatives and use skills, techniques, processes, standards and conventions to achieve information solutions.</td>
<td>Items from Character Design (Task 2) and Character Development (Task 3) and Healthy Lifestyle Promotion Task 4B.</td>
</tr>
<tr>
<td>3.1 Apply accepted standards and conventions to create or modify various forms of information solutions.</td>
<td>Items from Character Design (Task 2) and Character Development (Task 3) and Healthy Lifestyle Promotion Task 4B.</td>
</tr>
<tr>
<td>3.2 Apply efficient skills, techniques and processes in the use of ICT systems to develop information solutions.</td>
<td>Items from Character Design (Task 2) and Character Development (Task 3) and Healthy Lifestyle Promotion Task 4B.</td>
</tr>
<tr>
<td>3.3 Apply enterprising capabilities while exploring alternatives and working to achieve information solutions.</td>
<td>Items from Character Design (Task 2) and Character Development (Task 3) and Healthy Lifestyle Promotion Task 4B.</td>
</tr>
<tr>
<td>4 - Information and Communication Technologies in Society: Students understand how cultural beliefs, values, abilities and ethical positions are interconnected in the development and use of information and communication technologies</td>
<td>Items from Research Report (Task 1) and Attitudes test (Task 4A). Annotation for O4.</td>
</tr>
<tr>
<td>4.1 understand the cultural beliefs, values, abilities and ethical position that can impact on the use of ICT</td>
<td>Items from Research Report (Task 1) and Attitudes test (Task 4A). Annotation for O4.</td>
</tr>
<tr>
<td>4.2 understand the benefits and consequences of ICT use in different contexts and how this relates to beliefs, abilities, values and ethical positions</td>
<td>Items from Research Report (Task 1) and Attitudes test (Task 4A). Annotation for O4.</td>
</tr>
<tr>
<td>4.3 understand the consequences of technological developments on social, personal, cultural, physical, economic and ethical structures and environments</td>
<td>Items from Research Report (Task 1) and Attitudes test (Task 4A). Annotation for O4.</td>
</tr>
</tbody>
</table>
The two external assessors marked the student work on the aspects of the four outcomes (1, 2, 3, & 4) while the teacher allocated a level for each outcome and an overall level. The assessor judgements on the four outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The numeric scores for each assessor were summed across the four outcomes to produce a score with a maximum of 100 for each assessor. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Consensus (C)</th>
<th>Teacher (%)</th>
<th>Rank A1</th>
<th>Rank A2</th>
<th>Rank C</th>
<th>Rank Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>WA101 etc.</td>
<td>38</td>
<td>26</td>
<td>25</td>
<td>100</td>
<td>2</td>
<td>2.5</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

**Figure 4.2: Tabulation of results from marking.**

A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

**Correlations of Rankings**

<table>
<thead>
<tr>
<th>N=10</th>
<th>Rank of TOT_M1</th>
<th>Rank of TOT_M2</th>
<th>Rank of CONSENSUS</th>
<th>Rank of Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank of TOT_M1</td>
<td>Pearson Correlation</td>
<td>.938(**)</td>
<td>.925(**)</td>
<td>.718(*)</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.019</td>
<td></td>
</tr>
<tr>
<td>Rank of TOT_M2</td>
<td>Pearson Correlation</td>
<td>.938(**)</td>
<td>1</td>
<td>.603</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.065</td>
<td></td>
</tr>
<tr>
<td>Rank of CONSENSUS</td>
<td>Pearson Correlation</td>
<td>.925(**)</td>
<td>.997(**)</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.578</td>
<td></td>
</tr>
<tr>
<td>Rank of Tch</td>
<td>Pearson Correlation</td>
<td>.718(*)</td>
<td>.603</td>
<td>.578</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.019</td>
<td>.065</td>
<td>.080</td>
<td></td>
</tr>
</tbody>
</table>

**Correlations of Marks**

<table>
<thead>
<tr>
<th>N=10</th>
<th>TOT_M1</th>
<th>TOT_M2</th>
<th>CONSENSUS</th>
<th>Tch %</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT_M1</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.873(**)</td>
<td>.879(**)</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.001</td>
<td>.001</td>
<td>.067</td>
<td></td>
</tr>
<tr>
<td>TOT_M2</td>
<td>Pearson Correlation</td>
<td>.873(**)</td>
<td>1</td>
<td>.972(**)</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.001</td>
<td>.000</td>
<td>.134</td>
<td></td>
</tr>
<tr>
<td>CONSENSUS</td>
<td>Pearson Correlation</td>
<td>.879(**)</td>
<td>.972(**)</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.001</td>
<td>.000</td>
<td>.144</td>
<td></td>
</tr>
<tr>
<td>Tch</td>
<td>Pearson Correlation</td>
<td>.600</td>
<td>.508</td>
<td>.496</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.067</td>
<td>.134</td>
<td>.144</td>
<td></td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

**Correlation is significant at the 0.05 level (2-tailed).**

**Figure 4.3: Correlations between results from marking WA student work.**

The two external markers used the level statements for the four outcomes to mark the digital portfolio of work. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.938 (p<0.001) on the ranking of students.

Initially the external markers disagreed substantially on only three (WA101, WA104 and WA109) of the 10 students. Note that the difference in results for student WA101 and WA104 had little effect on the ranking. Four students (WA103, WA 105, WA106 and WA107) obtained an award of N/A (0) indicating they were below the lowest award of Level 4 (12) for all the outcomes. After a meeting between the external markers a consensus mark was allocated to each student with some time spent discussing the three students WA101, WA104 and WA109.
The teacher allocated a level for each outcome and also an overall level. The level ranged from 2 to 4, 4 being the highest awarded (20 marks were awarded for level 4, 15 marks for level 3, and 10 marks for level 2). The marks were then summed in order to obtain a mark out of 100 for each student, which could then be ranked.

Comparison with external marker results indicated limited agreement with the teacher’s results, with only four students ranked in similar positions. When compared with the two assessors the inter-rater reliability correlations on the ranking of students were 0.72 and 0.60 respectively, with the first assessor being significant at the 0.05 level of significance. There was very low correlation between the actual marks generated by the external markers and the teacher.

**Conclusions**

Overall it was concluded that this digital portfolio of work was implemented with a number of significant logistical difficulties with most students finding they needed more time to understand how the web editing software worked and a number of students did not have enough time to either complete the Five tasks or the digital portfolio within the time permitted.

Capturing the performance of the students was relatively easy because each student’s work files were provided in folders on a single DVD with a folder for each student. After checking the technical operation of the website each student’s folder was zipped as a web folder and uploading into the MAPS portfolio system. It could then be viewed directly through a browser by each marker.

In terms of the marking of the student work the following conclusions could be drawn.

1. There were clear differences in results based on the external marking and teacher marking by the outcomes using the standards framework. This was likely to be due to the teacher’s knowledge of the day-to-day work of the students that was not necessarily represented in their portfolios.

2. There was a high level of agreement between the two external assessors using the standards framework. Disagreements tended to relate to the interpretation of a student’s work when little supporting documentation was provided.

3. Only one student was judged to have demonstrated any work above Level 4 by the external markers. This agreed with the teacher’s judgement of the class as a whole.
**AA: Digital Portfolio Slideshow**

The AA case study involved one class of Year 11 students at one school completing the Applied Information Technology course Unit 1A in a Digital Graphics context. There were 17 students in the class but two students were not involved in the study, because either they did not want to be involved or their parents did not give permission.

The aim was for each student to collate a portfolio of work stored in digital formats and emanating from work completed in their course. The work related to the four outcomes for the course with items drawn from the two tasks completed during the semester.

The portfolio was developed around the first two sets of tasks from the West One package, *Using Information Technology*. The researcher discussed with the teacher the best vehicle for the presentation of the portfolios and it was decided to use Powerpoint as most students had some experience in using this software.

Students were required to construct their portfolio using items from the tasks they had completed throughout the course. For each task there was an indication of what to include in the portfolio.

Students were required to collate and display their portfolio as a slideshow. Most of the collation occurred in the last few weeks of Term 2. The teacher invigilated.

**Instructions for the Portfolio**

You will complete and submit an electronic portfolio. The portfolio will be used as evidence that you can edit, create and integrate interactive multimedia.

You are required to select digital artefacts and process documents that demonstrate the highest achievement in the creative and conceptual realisation of the application of information and communications technologies (ICT) to address human needs. The emphasis is on quality not quantity.

It contributes to the assessment of the student’s ability to use and evaluate the application of ICT relevant to personal, community, business and industry settings. The portfolio provides evidence for experience in which they have developed an understanding of, and practical skills in the use of ICT relevant to these settings.

Each digital file should include an annotation indicating how it was created, the format of the file, and an explanation of the purpose and uses of the file.

Students included a list of items from two tasks completed through the semester as indicated. These were from the West One package ‘Using Information Technology’. Create a slideshow to connect all the items and keep a web-log of your activities.

**Task 1: Event Package**

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Event Logo</td>
<td>Research and produce a logo for the event, keeping in mind the type of event and the intended audience, using a word processing or a desktop publishing application.</td>
</tr>
<tr>
<td>1.2</td>
<td>Running Sheet/Memo</td>
<td>Produce a detailed schedule which is called a running-sheet, indicating what is expected to happen during the event, when and who is involved. Use a word processing or spreadsheet application.</td>
</tr>
<tr>
<td>1.3</td>
<td>Ticket or Invitation</td>
<td>Produce a ticket or an invitation, whichever is more suitable, using a word processing, desktop publishing or presentation software.</td>
</tr>
<tr>
<td>1.4</td>
<td>Advertisement</td>
<td>Produce an advertisement, paper-based such as a brochure or flyer or a digital presentation.</td>
</tr>
<tr>
<td>1.5</td>
<td>Database of Suppliers</td>
<td>Develop a simple database containing the names and contact details of relevant suppliers from your area.</td>
</tr>
<tr>
<td>1.6</td>
<td>Mail Merge Letters</td>
<td>Set up a letter template to obtain quotes from suppliers.</td>
</tr>
<tr>
<td>1.7</td>
<td>Task 1 Evaluation</td>
<td>Fill in evaluation sheet</td>
</tr>
</tbody>
</table>
Task 2: Report on Proposal to buy a Home Office System

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Report Heading</td>
<td>A statement indicating who produced the report, which group it was produced for and why.</td>
</tr>
<tr>
<td>2.2</td>
<td>Introduction</td>
<td>Explain the importance of a home computer system to increases in efficiency and quality of work.</td>
</tr>
<tr>
<td>2.3</td>
<td>Hardware/Software</td>
<td>Describe each of the components typically used in a home computer system, both hardware and software, and uses of each.</td>
</tr>
<tr>
<td>2.4</td>
<td>Budget</td>
<td>Prepare a budget for the components of a home computer system. Include three choices using appropriate software (e.g. spreadsheet).</td>
</tr>
<tr>
<td>2.5</td>
<td>Skills Survey</td>
<td>Conduct and analyse a survey of individuals using a home computer system (who uses it and for what purposes).</td>
</tr>
<tr>
<td>2.6</td>
<td>Ergonomic Factors</td>
<td>Give an explanation of ergonomic factors that need to be considered when setting up and using a home computer system.</td>
</tr>
<tr>
<td>2.7</td>
<td>Future Trends</td>
<td>Comment on existing technologies and future trends in home computer system requirements.</td>
</tr>
<tr>
<td>2.8</td>
<td>Recommendations</td>
<td>Make a recommendation of the best system to purchase or upgrade to for the money, and on its suitability for the home environment.</td>
</tr>
<tr>
<td>2.9</td>
<td>Task 2 Evaluation</td>
<td>Fill in evaluation sheet</td>
</tr>
</tbody>
</table>

Results

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class

The class was observed for one session in June while they were working on their portfolios. The students worked in a laboratory containing WinXP workstations.

The class of 10 students (3 others were absent) did the survey first and then continued working on their portfolios that had been a focus over the last 2 weeks. This was based around the West One package of tasks “Using Information Technology” – these instructions had been uploaded to the school Intranet. Most were using Powerpoint to collate the portfolios by inserting content from folders organised by weeks or linking to files in these folders. Two students said they had finished, 1 was emailing, 2 were surfing the web, and 1 was looking at the instructions online. Two students said they didn’t want to be in the study because they didn’t like people looking at their work. Towards the end one student was doing an evaluation document and one was looking at Google Earth/Maps. Two students who were finished started on unit 1B doing the “Turn It Up” module.

Survey of Students

There were 10 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.
Table 4.2
Descriptions and descriptive statistics for the scales based on items from the student questionnaire (N=10).

<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>10</td>
<td>2.4</td>
<td>3.4</td>
<td>2.93</td>
<td>0.28</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>10</td>
<td>1.4</td>
<td>3.0</td>
<td>2.14</td>
<td>0.49</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>10</td>
<td>1.6</td>
<td>2.8</td>
<td>2.26</td>
<td>0.34</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confidence</td>
<td>10</td>
<td>2.2</td>
<td>2.7</td>
<td>2.40</td>
<td>0.18</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>10</td>
<td>2.7</td>
<td>3.9</td>
<td>3.38</td>
<td>0.38</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>10</td>
<td>50</td>
<td>156</td>
<td>86.9</td>
<td>31.7</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

Figure 4.4: Distribution of scores for the Skills and SCuse scales on the student questionnaire.

From the table above it is clear that the students didn’t tend to have done the nominated activities in Q10 and about half were not confident they could do them except by typing an assignment for school that they all had done. Most liked using computers at school and home but only sometimes liked finding out things for themselves. Six of the ten indicated great confidence in using computers. Eight had Broadband Internet access at home and nine used computers on a daily basis.

Students indicated that the assessment was relatively easy to complete, that they typically had not done the range of applications indicated, that they were fairly confident in using ICT, and had a relatively high level of ICT skills across the range of types of applications. On average they indicated using ICT for about one and a half hours per day at school with the lowest 50 minutes and the highest 156 minutes but the majority tended to be below the mean.

Almost all Agreed or Strongly Agreed with Q2 items concerning value of eAssessment (only 2 disagreed on any of these items). Half indicated having little or no experience with electronic portfolios and felt that it would take a while to get used to it.

There were two open-ended questions on the student questionnaire. Item 3 asked students to list the two best things about the assessment and Item 4 asked them to list the two worst things about the assessment. Generally students considered that using computers made it easier, quicker and more interesting than using pen and paper. Three students appeared quite negative and probably did not like doing any work on computers. Four students provided no ‘worst things’.


**Students Forum**

The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.

The students indicated that while the tasks were time consuming they did help to develop IT skills. They believed that the more effort they put in they produced better quality work. All but one student had not done a digital portfolio before, with all previous portfolios done on paper. They identified a few technical problems with the server crashing a few times that led to a loss to time. And some difficulties in hyperlinking from Powerpoint to PDFs and CorelDraw files. They would like choice of content topics that would lead to increased enjoyment and would like to learn more about software and how to use it. Further, they would like less week-by-week tasks and reduce the time constraints, that is, they would like more open-ended longer tasks. However, it was better than doing work on paper.

**Structured Email ‘Interview’ with Teacher**

The teacher believed that the structure for the portfolio suited the better students but that a more formal structure was needed for the lower ability students in this Stage 1 unit. He felt that those students who set up their digital portfolio as they went found the task easy to complete and were therefore positive while a few others reacted, "Oh well another task". Also the better students had the IT skills to quickly adapt what they had already produced for inclusion in the task. So the better students complied readily and although they didn't like all the requirements methodically completed the task.

The teacher said that in general the students don’t mind a digital presentation if it is a chance to show off their skills and innovative solutions and some are also aware of the loss of "wow" factor that can occur when a hard copy is presented. However, one student did not mind the teacher viewing a digital copy but did not like others to see it. He believed that generally the work submitted was of a good standard and lower grades were mainly a result of omission rather than low quality.

There were no technical problems although for some students a lack of IT skills and understanding of file structure led to "lost" files. In some cases the quality of digital images was lacking because the cameras used were low end due to financial constraints. He believed that there was great potential for using electronic portfolios but that this required very specific assessment criteria and pointers to enable an assessment of level attained (a rubric would be essential). He felt that many teachers with weak IT backgrounds would need a lot of PD and perhaps a "How to" was required for both staff and students. It would be useful to have specialist portfolio creation software with the collation ongoing not just as the end of a course task. He felt that some students had a facade of "I already know this and don't need instruction/help" to set up the filing structures required to implement a digital portfolio but that they had only superficial skills and therefore needed a good example of a setup.

**Results of Marking**

The portfolio output for 15 students were collected on a DVD, copied from the school’s server upon which the students saved their work. Each student had a folder containing folders and files that made up their portfolio. The formats of the files included HTML, DOC, PUB, and PPT. Before marking could commence all the portfolios were checked, zipped and uploaded into the MAPS system.

The students’ work was marked by two external assessors using a standards-based rubric based on criteria associated with the four outcomes of the course, and by the teacher awarding a level for each outcome. For the standards-based marking method the assessors used the rubric provided in the course document based on the four outcomes, as described below. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the instructions in the table below on the criteria to assess. The portfolio was based on two tasks and intended to be presented through a slideshow. Task 1 was titled, Stationery for an Event, and consisted of a producing a set of office documents. Task 2 was a word-processed report on the ‘Selection of Computer System’ for a business.
<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Technology Process: Students apply a technology process when creating or</td>
<td></td>
</tr>
<tr>
<td>modifying products using information communication technologies.</td>
<td></td>
</tr>
<tr>
<td>1.1 investigate ideas, and evaluate alternative designs and proposals</td>
<td>Task 1 – Range of alternatives for stationery products and explanation of development of final products.</td>
</tr>
<tr>
<td></td>
<td>Task 2 – Investigates at least two PC systems and office furniture. Sources of information, budget and explanation of choices in terms of needs of clients. Survey to determine client skills.</td>
</tr>
<tr>
<td>1.2 devise communicate and evaluate proposals and design plans in appropriate</td>
<td>Task 1 – Methods of communicating design alternatives for stationery products (e.g. sketches and mock-ups) and explanation of choices (e.g. cost, ergonomics, clients needs).</td>
</tr>
<tr>
<td>forms</td>
<td>Task 2 – NO INFORMATION.</td>
</tr>
<tr>
<td>1.3 implement and evaluate production processes and strategies to manage</td>
<td>NO INFORMATION</td>
</tr>
<tr>
<td>resources efficiently</td>
<td></td>
</tr>
<tr>
<td>2 - Understanding Information and Communication Technologies: Students</td>
<td></td>
</tr>
<tr>
<td>understand the nature and use of computer hardware and software to achieve</td>
<td></td>
</tr>
<tr>
<td>information solutions.</td>
<td></td>
</tr>
<tr>
<td>2.1 understand the ICT-related concepts, standards and terminology required</td>
<td>Task 1 – NO INFORMATION</td>
</tr>
<tr>
<td>to select and use appropriate computer software and hardware to achieve</td>
<td>Task 2 – Explanation of choices shows understanding of ICT systems and uses appropriate terminology.</td>
</tr>
<tr>
<td>information solutions</td>
<td></td>
</tr>
<tr>
<td>2.2 understand the standards and conventions when selecting and using ICT</td>
<td>Task 1 – NO INFORMATION</td>
</tr>
<tr>
<td>systems to achieve information solutions</td>
<td>Task 2 – Explanation of choices shows understanding of standards and conventions of ICT systems.</td>
</tr>
<tr>
<td>2.3 understand the management, processes, functions, types and relationships</td>
<td>NO INFORMATION</td>
</tr>
<tr>
<td>between the components of ICT systems required to achieve information</td>
<td></td>
</tr>
<tr>
<td>solutions</td>
<td></td>
</tr>
<tr>
<td>3 - Quality of Information Solutions: Students explore alternatives and use</td>
<td></td>
</tr>
<tr>
<td>skills, techniques, processes, standards and conventions to achieve</td>
<td></td>
</tr>
<tr>
<td>information solutions.</td>
<td></td>
</tr>
<tr>
<td>3.1 Apply accepted standards and conventions to create or modify various</td>
<td>Task 1 – Applies standards and conventions for business stationery (logos, letters, posters) such as positioning, simple graphics, consistency of look and feel.</td>
</tr>
<tr>
<td>forms of information solutions.</td>
<td>Task 2 – NO INFORMATION</td>
</tr>
<tr>
<td>3.2 Apply efficient skills, techniques and processes in the use of ICT</td>
<td></td>
</tr>
<tr>
<td>systems to develop information solutions</td>
<td></td>
</tr>
<tr>
<td>3.3 Apply enterprising capabilities while exploring alternatives and working</td>
<td>Task 1 – Some innovative approach to design of business stationery.</td>
</tr>
<tr>
<td>to achieve information solutions.</td>
<td>Task 2 – NO INFORMATION</td>
</tr>
<tr>
<td>4 - Information and Communication Technologies in Society: Students understand</td>
<td></td>
</tr>
<tr>
<td>how cultural beliefs, values, abilities and ethical positions are</td>
<td></td>
</tr>
<tr>
<td>interconnected in the development and use of information and communication</td>
<td></td>
</tr>
<tr>
<td>technologies</td>
<td></td>
</tr>
<tr>
<td>4.1 understand the cultural beliefs, values, abilities and ethical position</td>
<td>NOT ENOUGH TO ASSESS ANY STUDENTS</td>
</tr>
<tr>
<td>that can impact on the use of ICT</td>
<td></td>
</tr>
<tr>
<td>4.2 understand the benefits and consequences of ICT use in different contexts</td>
<td>NOT ENOUGH TO ASSESS ANY STUDENTS</td>
</tr>
<tr>
<td>and how this relates to beliefs, abilities, values and ethical positions</td>
<td></td>
</tr>
<tr>
<td>4.3 understand the consequences of technological developments on social</td>
<td>NOT ENOUGH TO ASSESS ANY STUDENTS</td>
</tr>
<tr>
<td>personal, cultural, physical, economic and ethical structures and environments</td>
<td></td>
</tr>
</tbody>
</table>
The two external assessors marked the student work on the aspects of the four outcomes while the teacher marked also using the four outcomes. The assessor judgements on the outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The numeric scores for each assessor were summed across the outcomes to produce a score with a maximum of 100 for each assessor. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Consensus (C)</th>
<th>Teacher Level</th>
<th>Teacher %</th>
<th>Rank A1</th>
<th>Rank A2</th>
<th>Rank C</th>
<th>Tch Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA101</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>C</td>
<td>61</td>
<td>6</td>
<td>11.5</td>
<td>7</td>
<td>7.5</td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 4.5: Tabulation of results from marking.*

A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

*Figure 4.6: Correlations between results from marking AA student work.*

Two external markers used the level statements for the outcomes (1 and 3) to mark the portfolios. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.87 (p<0.001) on the ranking of students. This increased with some discussion of the results for students 104, 109 and 114. In part the high correlation was due to the large number of portfolios allocated a score of 0 (being judged to be below Level 4). The main discrepancies between the two markers
concerned interpretations of the degree to which students had linked research, the characteristics of audience, and design principles to the development of their design.

The teacher provided a percentage mark and grade for the portfolio and a grade for the semester. Comparison with external marker results indicated reasonable agreement on ranking except perhaps student 106. When compared with the two assessors and the consensus the inter-rater reliability correlations on the ranking of students were 0.63, 0.58 and 0.63 respectively, with all significant at the 0.05 level of significance. It should be noted that the external assessors could only mark on the material provided in the digital portfolio that may not necessarily reflect all the work the student had done that the teacher would be aware of when marking.

Conclusions

Overall it was concluded that this electronic portfolio assessment was implemented with no significant technical difficulties with only lack of student skill and experience constraining the completion and quality of work submitted. The use of the West One package of tasks “Using Information Technology” provided a high degree of scaffolding that it appeared most of the students needed, although some complained that there was not enough freedom to develop their own ideas.

Capturing the performance of the students was somewhat complex with student work provided in individual folders on a CDROM. For most students work was organised into weekly folders. In some cases work was also organised in Task folders and in some cases also a Portfolio folder. The completeness of the portfolio shows to access the portfolio items varied considerably and was not adequate to ensure markers had access to all portfolio items. Therefore most students’ portfolios needed further work to gather all portfolio items into a separate folder for each student and files renamed by Task Item number before zipping the folder and uploading this into the MAPS system.

The structure of the two tasks limited the opportunity for students to demonstrate more than a L5 of achievement and difficult for assessors to recognise that achievement. However, it probably ensured that at least half could demonstrate at least a L4. It was likely that many of the other students were unlikely to achieve this level at this time in any activities.

In terms of the marking of the student work the following conclusions could be drawn.

1. There was only adequate student work to assess Outcomes 1 and 3 and for many students this was not even possible. In particular most students lacked sufficient process documentation to make judgements for Outcome 1 and annotation to make judgements for Outcome 3.
2. There was little difference between the assessors and between the assessors and teachers, in the ranking of student work. This was in part due to the high number of students who demonstrated achievement below Level 4.
3. There was a high level of agreement between the two external assessors using the standards framework although there were differences with a few students in relation to the interpretation of the degree to which their designs were intentional.
4. Three students were initially judged to have demonstrated work above L4 by only one assessor.
5. For these students in a Stage 1 unit the task of compiling a portfolio as a slideshow appeared to complex and it may have been best to collate the items into TWO task documents (Word or PDF).

Feedback from Teacher

I found the conclusions very interesting and will certainly change some of my processes in presenting the course as a result of the project. I shall also concentrate on file management skills as well as upgrading their skills in applications such as PowerPoint (which most said that they could use but were limited in the use of a number of attributes available). I concur with the findings of the report and found that for AIT 1B, a similar pattern of results occurred although there was considerable improvement in portfolio development by most of the students. Unfortunately some still failed to complete set work with time constraints the general excuse. Your observation regarding the productivity of a few of the students during class time prevailed and they looked at the final dates for submission of the tasks rather than setting short-term goals. Hence, the feeling that there is plenty of time to finish this task, led to final panic and then no product to submit.
IA: Electronic Portfolio using Flash

The IA case study involved one class of Year 12 students at one school completing the Interactive Media course. There were 13 students in the class.

The aim was for each student to collate a portfolio of work stored in digital formats and emanating from work completed in their course. The work related to the outcomes being assessed and conformed to given parameters for submission of the portfolio – form, structure, range of samples required.

The teacher developed the portfolio format and instructions based on approaches used in the previous few years. This had already been incorporated into the programme for the year and therefore no changes were made based on discussions between the teacher and researcher. Students were provided with instructions about the portfolio at the beginning of the year.

Instructions for the Portfolio

Students completed and submitted an electronic portfolio. The portfolio was to be used at the end of the year as evidence that the students could edit, create and integrate interactive multimedia. This task required that each student create a portfolio of interactive multimedia, which was his/her own work, together with notes explaining how each piece of work was created. The students were given the following information on what the portfolio was to include.

1. Files in the portfolio need to demonstrate the following features:
   - Navigation systems (e.g. book metaphor, console, hierarchical, map/desktop)
   - Interaction methods (e.g. menu choices, text entry, click touch areas)
   - Animations (Motion and/or shape tweening created in Flash)
   - Audio output (Edited and Original samples of mp3 music)
   - Video (Edited to include Graphics, Video, Audio and Text)

2. The portfolio must utilise a logical navigation structure.
   - This could either be done by the tasks you completed this year or by relevant categories.
   - You need to utilise an interface to ease navigation.
   - A help page explaining how to use your portfolio could also be included.

3. Notes are to be included to explain how the individual items were put together.
   - These are to be contained within the presentation and must include:
     (i) A list of software applications that were used to create the items,
     (ii) The time spent on the project,
     (iii) A break down of the individual skills that you learned.
     - Using headings such as “Outline”, “Software Used”, “Time” and “Skills Learned”.

4. The portfolio is to be in the form of a Macromedia Flash presentation burned onto a CD.
   - A CD cover, label and back index need to be designed.

The teacher assessed the portfolio on four outcomes in the Interactive Media subject using the criteria shown in the following table.
# Assessment Grid

<table>
<thead>
<tr>
<th>Task</th>
<th>Task Description</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Five</td>
<td>Digital Portfolio</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requirements</th>
<th>OUTCOME</th>
<th>EX</th>
<th>VG</th>
<th>G</th>
<th>OK</th>
<th>POOR</th>
<th>MISSING</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Work Ethic</strong></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focus on Task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisation Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meeting Deadlines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Technology Process</strong></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self Critique</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contingency Allowance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisation of class time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initiative and responsibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Software Skills</strong></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphic Manipulation/Photoshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audio Manipulation/Audacity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animation Manipulation/Flash</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimedia Programming/MovieMaker2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Development/Dreamweaver</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appropriateness of Software</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Portfolio Features</strong></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Impact</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interface/Button Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ease of Navigation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of Animation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text Legibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphic Display</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audio Output</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colour Scheme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creative/Unusual Elements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**General Comments:**

- **Work Ethic:**
- **Technology Process:**
- **Software Skills:**
- **Portfolio Features:**
Results
A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class
Towards the end of the first semester the 13 students were observed working on the portfolio in a computer laboratory for one hour. The class had been working on the portfolios for last two weeks using Flash on HP workstations (WinXP) to collate files from the first four tasks of the Interactive Media course and include annotations connected to the outcomes of the course. Files were imported or linked to a main Flash file that was the vehicle for an interface based on a student generated metaphor.

Initially most students were working on the interface with some attaching work. All were very focussed on the work sitting at two rows of computers (back to back). They worked independently but called on the teacher and each other for help. For example, the teacher showed a student how to link Flash files. Some students were using Photoshop to edit graphics using Google images (no referencing evident). Two students had a problem with video import – to embed or stream the video (tried Help file and then asked teacher). Some were using headphones.

Typically the teacher made suggestions but it was the student’s choice what to do (for one student the teacher was more persuasive leading to the student following the teacher’s suggestion regarding an interface design). Some students looked at a syllabus statement for the course for portfolio task. Students helped each other but each did own work. Each student’s interface was very different. Most used rollovers for buttons on the interface. Some scripting of buttons was included. One student’s button didn’t work so the teacher worked on it to find the problem in the script.

After 50 mins the intensity of work and focus reduced for a few but all were still working well. All seemed to enjoy the creative opportunity and had a good understanding of Flash. Near the end one student had problems saving work and lost about 30 mins work. One student used Moviemaker. One student sent a quick personal email.

Survey of Students
There were 11 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below with scores calculated so that high scores are more positive.

Table 4.3

<table>
<thead>
<tr>
<th>Description</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess: Ease of assessment for the students</td>
<td>11</td>
<td>2.7</td>
<td>3.6</td>
<td>3.2</td>
<td>0.25</td>
</tr>
<tr>
<td>Apply: Application of some computer uses</td>
<td>11</td>
<td>2.2</td>
<td>3.0</td>
<td>2.6</td>
<td>0.26</td>
</tr>
<tr>
<td>Attitude: Attitude towards using computers</td>
<td>11</td>
<td>1.8</td>
<td>2.6</td>
<td>2.4</td>
<td>0.22</td>
</tr>
<tr>
<td>Confidence: Confidence in using computers</td>
<td>11</td>
<td>2.2</td>
<td>2.7</td>
<td>2.5</td>
<td>0.15</td>
</tr>
<tr>
<td>Skills: A measure of ICT skills</td>
<td>11</td>
<td>2.9</td>
<td>3.7</td>
<td>3.4</td>
<td>0.27</td>
</tr>
<tr>
<td>SCuse: Estimate of time per day spent using computers at school</td>
<td>11</td>
<td>40</td>
<td>264</td>
<td>112</td>
<td>66.0</td>
</tr>
</tbody>
</table>
From the table above it is clear that the students had tended to have done the nominated activities in Q10 or were confident they could do them. All liked using computers at school and most at home but only sometimes liked finding out things for themselves. All indicated great confidence in using computers. All had broadband Internet access at home and eight of the 11 used computers on a daily basis at home for schoolwork.

Almost all Agreed or Strongly Agreed with Q2 items concerning value of eAssessment (only 2 disagreed on any of these items – B and H). Most had little experience with portfolios on computers with six indicating that they had never done it before but felt that it only took a little time to get used to it.

Students indicated that the assessment was relatively easy to complete, that they typically had done the range of applications indicated, that they were confident in using ICT, and had a high level of ICT skills across the range of types of applications. On average they indicated using ICT for nearly two hours per day at school with the lowest 40 minutes and the highest 264 minutes but the majority tended to be below the mean. Students indicated a high level of ICT skills across the range of applications. On average they indicated using ICT for just under two hours per day at school with the lowest 40 minutes and the highest 264 minutes.

There were two open-ended questions on the student questionnaire. Item 3 asked students to list the two best things about the assessment and Item 4 asked them to list the two worst things about the assessment. Generally students considered that using computers made constructing the portfolio more interesting, fun, easier, and allowed them to develop new skills. They were concerned with the amount of time it took to create the portfolios, some found it difficult and some were concerned about the reliability of the hardware or software and the possibility of losing work.

**Students Forum**

The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.

The students explained that they had done quite a few tasks culminating in the digital portfolio. They had edited a video file and video animation, changed a music file, done an essay and done a test and put it all in a digital portfolio in Flash. They thought constructing the portfolio was ‘pretty good’, such as how to make buttons and put everything together in one file to allow it to be easily viewed. They had done some of this last year but not a digital portfolio like this. They generally felt that most students liked it and found it is interesting as they could use skills like to make movie stuff and
learning things like editing. They were still learning options to do with PowerPoint for students to advance their knowledge. They started at the beginning of the year, and do tasks then put into the portfolio.

They felt that they could ‘sort of’ demonstrate the quality of their work because they could put in all personal things like art; they could have choice about what they learned. They felt that it separated the students who really liked learning and those who just sit and struggle and don’t do anything. They believed it was different because in the past everything was marked separately, printed out and kept in a file separately. Now they could go back to what you have done before, review all tasks or skills you learnt which was better, especially when they put in graphics and things, previously they might only do a little bit. They found the Flash programming difficult, particularly because they did not have the program at home so they had to do it all during the school time. They found that sometimes the program “stuffs up” at school, particularly with large file sizes.

They could not identify any technical problems except things like making a link. They had their own network space and we could access this from any computer and also at home if they had the software. They could also save school files on a thumb drive. They could not identify any other significant problems. They felt that they all worked well as a team, helping each other and because they had had the whole year they didn’t have time stress. However, they thought that it would be better if for some graphics they went out and took photos and videos and made it more professional. They could do it now but thought it would be too time consuming so they just drew pictures.

**Structured Email ‘Interview’ with Teacher**

The teacher believed that the task suited both his needs as a genuine assessment tool and the interests of the students as a “professional looking showcase” of manipulating multimedia, and as such was an integral part of the course. He definitely felt that it provided a more “accurate picture of the students specific abilities than an ‘essay-type’ of assessment”. He felt that the processes had worked well and there were no significant problems. He was pleased with the quality of the students’ work and felt that the performance of students was as expected with higher skill levels transferring to higher quality portfolios.

He was confident that the students enjoyed doing the portfolio but that it required a high level of IT skill. As such this type of assessment was only relevant to “computer based subjects”. He emphasised that the portfolio needed to be ongoing with the students coming back regularly. This was because they develop skills throughout the year and it is better to include work immediately after it has been completed so that they remember what it was about.

**Results of Marking**

The portfolio output for 13 students were collected on a DVD, copied from the school’s server upon which the students saved their work. Each student had a folder containing folders and files that made up their portfolio. The formats of the files included FLV, HTML, DOC, PUB, MOV and PPT. Before marking could commence all the portfolios were checked, zipped and uploaded into the MAPS online.

The students’ work was marked by two external assessors using a standards-based rubric based on criteria associated with outcomes 1 and 3 of the course, and by the teacher awarding a level for each of four outcomes of the Interactive Media subject. For the standards-based marking method the assessors used the rubric provided in the course document based on the four outcomes, as described below. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the instructions in the table below on the criteria to assess. The portfolio was created in Flash with the student’s own design for the interface, including a range of work with explanations. This was collated towards the end of Semester One and would be further updated towards the end of Semester Two. The portfolio should include: a logical navigation system based on a metaphor; a variety of interaction methods; animation; audio; video; notes to explain how the individual items were put together; a digital CV; and collated as a Macromedia Flash presentation.
<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Technology Process: Students apply a technology process when creating or modifying products using information communication technologies.</td>
<td>Portfolio items – Investigations/research conducted. Considers alternatives in designs.</td>
</tr>
<tr>
<td>1.1 investigate ideas, and evaluate alternative designs and proposals</td>
<td>Portfolio items – Quality of design plans. Forms of communication used (e.g. storyboards, tables, descriptions).</td>
</tr>
<tr>
<td>1.2 devise communicate and evaluate proposals and design plans in appropriate forms</td>
<td>Portfolio items – Quality of discussion in annotations about production and evaluation.</td>
</tr>
<tr>
<td>1.3 implement and evaluate production processes and strategies to manage resources efficiently</td>
<td>Portfolio items – Application of IM standards and conventions in design of portfolio.</td>
</tr>
<tr>
<td></td>
<td>Portfolio items – Application of IM standards and conventions in design of the portfolio items themselves.</td>
</tr>
<tr>
<td></td>
<td>Portfolio annotations – Annotations show how applied in items.</td>
</tr>
<tr>
<td>3 - Quality of Information Solutions: Students explore alternatives and use skills, techniques, processes, standards and conventions to achieve information solutions.</td>
<td>Portfolio design – Application of IM skills and techniques in Flash for portfolio.</td>
</tr>
<tr>
<td>3.1 Apply accepted standards and conventions to create or modify various forms of information solutions.</td>
<td>Portfolio items – Application of IM skills and techniques in a range of software for portfolio items.</td>
</tr>
<tr>
<td></td>
<td>Portfolio annotations – Annotations show how applied in items.</td>
</tr>
<tr>
<td>3.2 Apply efficient skills, techniques and processes in the use of ICT systems to develop information solutions.</td>
<td>Portfolio design – Creativity and innovation in design of portfolio.</td>
</tr>
<tr>
<td></td>
<td>Portfolio items – Creativity and innovation in design and production of items in portfolio. Extent of creative design to cater for variety of users and purpose of information product.</td>
</tr>
<tr>
<td></td>
<td>Portfolio annotations – Explanation of design to cater for variety of users and purpose of information product.</td>
</tr>
<tr>
<td>3.3 Apply enterprising capabilities while exploring alternatives and working to achieve information solutions.</td>
<td>Portfolio design – Application of IM standards and conventions in design of portfolio.</td>
</tr>
<tr>
<td></td>
<td>Portfolio items – Application of IM standards and conventions in design of the portfolio items themselves.</td>
</tr>
<tr>
<td></td>
<td>Portfolio annotations – Annotations show how applied in items.</td>
</tr>
</tbody>
</table>

The teacher’s assessment addressed Outcomes 5, 6, 7 and 9 of the Interactive Media subject and used the related descriptors for Satisfactory, High and Very High performance.

Outcome 5: Works independently, collaboratively and in teams to achieve individual and shared goals within a realistic time frame.

Outcome 6: Describes and evaluates relevant information technology systems and process used in interactive media development.

Outcome 7: Uses appropriate software and equipment relevant to the design, production and marketing of interactive media products.

Outcome 9: Develops creative approaches towards the development of interactive media.

The two external assessors marked the student work on the aspects of the two outcomes (1 & 3), while the teacher marked using grade descriptors for four outcomes from the Interactive Media subject. To obtain an overall mark each grade was allocated a mark equivalent and then summed and multiplied by five. In this case V=5, H=4, S=2.5 and ND=0. Both markers found that for only one student was there adequate evidence to justify awarding a level of 4 or better for Outcome 1 and therefore only the results for Outcome 3 were used. The assessor judgements were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The numeric scores for each assessor produced a score with a maximum of 25 for each assessor. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.
A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

Two external markers used the level statements for Outcomes 1 and 3 to mark the portfolios. However, both markers found that for only one student was there adequate evidence to justify awarding a level of 4 or better for Outcome 1 and therefore only the results for Outcome 3 were used. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.92 ($p<0.001$) on the ranking of students. This increased with some discussion of the results for students 104 and 108. Discrepancies between the markers appeared to emanate from differences in interpreting the degree to which students had applied design principles to the user-interface design of the portfolio with only two students providing documentation of design decisions. There were also some differences in the understanding of the degree of difficulty in the Flash programming.
The teacher used the grade descriptors for four of the outcomes from Interactive Media (5, 6, 7 & 9). To obtain an overall mark each grade was allocated a mark equivalent and then summed and multiplied by five. In this case V=5, H=4, S=2.5 and ND=0.

Comparison with external marker results indicated some level of agreement given that two different sets of criteria were used. There were only substantial differences in ranking on five students. When compared with the two assessors and consensus the inter-rater reliability correlations on the ranking of students were 0.32, 0.32 and 0.34 respectively, with none significant at the 0.05 level of significance.

It should be noted that it is difficult to draw any conclusions when comparing the teacher marking with the external marking because different outcomes were assessed with the students actually completing one of the old subjects, Interactive Media. These outcomes were considerably different to those for the new course.

**Conclusions**

Overall it was concluded that this digital portfolio was implemented with little technical difficulties with all students able to complete the requirements of the portfolio to some extent. A few students indicated finding it difficult and a few lost substantial amounts of work at various times. A number of the portfolios submitted at the end of Semester One were inadequately collated to permit easy access to all their work and therefore the portfolios were submitted again towards the end of Term Three. The students responded well to the style of assessment despite their stated lack of experience. They felt it was interesting and they learned a lot from the activity. As a group of students they appeared particularly highly skilled, technically, and very confident in their expertise. They also appeared to use computers for a large proportion of time both at school and home.

Capturing the performance of the students was relatively complex because each student’s work files needed to be individually checked to ensure that all linked files were available and that buttons worked. The students’ work was provided in folders on a single DVD with a folder for each student. After checking the technical operation of the Flash presentation file and in some cases performing some maintenance each student’s folder was zipped as a web folder and uploading into the MAPS portfolio system. It could then be viewed directly through a browser by each marker.

The limited annotations and lack of documentation from most students made it difficult for students to demonstrate high level achievement and difficult for assessors to recognise that achievement. Further, since most of the portfolio items were class exercises each student’s capability could largely only be judged from the interface they had designed for the portfolio. However, only two students provided a document clearly explaining the design of their portfolio. A group project involving the design of a website was also included so this was used in marking as if each student had completed it individually.

In terms of the marking of the student work the following conclusions could be drawn.

1. There were clear differences in results between the two external assessors and the teacher because they were using two different sets of outcomes.
2. There was a high level of agreement between the two external assessors using the standards framework on Outcome 3. It was decided not to use Outcome 1 as only two students provided documentation relating to technology process.
3. Five students were judged to have demonstrated work above Level 4 by at least one assessor with one potential Level 6, although there was some disagreement due to differences in the interpretation of the application of design principles and the degree of difficulty in the Flash programming.
PA: Digital Portfolio with Format Choice

The PA case study involved two teachers and two classes of Year 11 students at one school completing the Applied Information Technology course Unit 2A in the Digital Graphics context. There were 20 students in total in the classes.

The aim was for each student to collate a portfolio of work stored in digital formats and emanating from work completed in their course. The work had to relate to the outcomes being assessed and conform to given parameters for submission of the portfolio – form, structure, range of samples required.

The teachers developed the portfolio based on the structure provided by the Curriculum Council. Detailed information about this was provided to the students at the beginning of the year along with details of all the other assessment tasks. As a result the researcher made no suggestions for changes to the portfolio.

Students were required to construct their portfolio using items from the tasks they had completed throughout the course. For each task there was an indication of what to include in the portfolio. Students were permitted to collate and display their portfolio as either a slideshow, website or Flash movie. Most of the collation occurred in the last few weeks of Term 2 with students permitted to work on their portfolios at home.

Instructions for the Portfolio

The following instructions were given to students as a guide to constructing their portfolio:

“When constructing an e portfolio, you need to consider what your personal theme will be. This is the colour choices and graphics that you will use through out your e portfolio. Remember that the colour wheels can help with colour selections.

How will you present your e portfolio? This is not intended to be prescriptive. You can choose whatever software you have access to when creating your e portfolio BUT with consideration for the person marking it. Remember, the only software you are sure they have is: Microsoft Office 2003; Internet Explorer 6; Adobe Acrobat Reader; and swf plugin for Internet Explorer.

You need to show the two best examples that you have in each of the aspects. These should not be entire tasks, rather the elements of the tasks that show you achievement in each aspect. There are 4 outcomes, each with 3 aspects and 2 examples. That is 24 examples of your work.”

Results

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the classes

Towards the end of the first semester the two classes of 9 and 11 students respectively did the survey first and then continued working on their portfolios that had been a focus over the last 2 weeks. They had choice of PowerPoint, a website using Frontpage, or Flash to use to collate and display their portfolio.

Class 1: Three students using Photoshop to create title graphics for portfolios. Four students using PowerPoint and tended to focus on ‘looks’ rather than content but one was developing ‘Outcome’ buttons for an interface. Two were looking at the ‘Instructions’ file and the teacher discussed these with one student and what type of work to put in. One student was using Google images, one coding HTML in Word, and one using Illustrator.

Class 2: Three students using Flash (they had done work on it last year), five using FrontPage and three using PowerPoint (tended to start from templates in FrontPage or PowerPoint). Also using Photoshop and Illustrator for title graphics. One student worked on a very detailed colour scheme (used # codes). Students helped each other with skills. Most working on an interface design for the
portfolio – very little content. One student scanned work to include. Two students used an online tool to create a “glowing” banner of text. Also some students used Google images. Teacher reminded about referencing graphics.

**Survey of Students**

There were 20 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below with scores calculated so that high scores are more positive.

**Table 4.4**

Descriptive statistics for the scales based on items from the student questionnaire (N=20).

<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>20</td>
<td>2.3</td>
<td>3.9</td>
<td>3.17</td>
<td>0.45</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>20</td>
<td>1.4</td>
<td>3.0</td>
<td>2.46</td>
<td>0.34</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>20</td>
<td>1.6</td>
<td>2.6</td>
<td>2.25</td>
<td>0.29</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confidence</td>
<td>20</td>
<td>1.3</td>
<td>2.8</td>
<td>2.32</td>
<td>0.44</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>20</td>
<td>2.0</td>
<td>3.6</td>
<td>3.21</td>
<td>0.45</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>20</td>
<td>6</td>
<td>192</td>
<td>75</td>
<td>47.9</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

**Figure 4.10: Distribution of scores for the Skills and SCuse scales on the student questionnaire.**

From the table above it is clear that most of the students had done the nominated activities in Q10 and were confident they could do them. Most liked using computers at school and home but only sometimes liked finding out things for themselves. Nineteen of the 20 indicated great confidence in using computers. Nineteen had Broadband Internet access at home and 17 used computers on a daily basis at home.

Students indicated that the assessment was relatively easy to complete, that they typically had done the range of applications indicated, that they were fairly confident in using ICT, and had a high level of ICT skills across the range of types of applications. On average they indicated using ICT for one and a quarter hours per day at school with the lowest 6 minutes and the highest 192 minutes (based on use at school in the previous week). However, most were below the mean.
Most Agreed or Strongly Agreed with Q2 items concerning value of eAssessment although six or seven did not agree that the steps of the portfolio were easy to follow or that the steps helped them develop their ideas. Thirteen indicated little or no experience with electronic portfolios although only six felt that it would take time to get used to it.

There were two open-ended questions on the student questionnaire. Item 3 asked students to list the two best things about the assessment and Item 4 asked them to list the two worst things about the assessment. Generally students considered that using computers made it easier, quicker and more professional looking. Some enjoyed the creation of the portfolio. Generally they found they did not have enough time, some were not sure what to do and a few indicated a lack confidence or skill. A few were concerned about the possibility of losing work.

**Students Forum**

The students interview forum was conducted by the researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.

The students said that they liked doing the portfolios and all had done digital portfolios before. Some of them had done a lot of graphics in Year 10 so found it relatively easy. Some chose to use Flash because they had done it in Year 10. Those who chose to present it as a website did so because it was easier to set up the links. Those who chose Powerpoint did so because it was easy. They believed that they were able to do their best quality work and none indicated experiencing any real problems in creating the portfolio. None of these students believed that anything needed changing and they could not recall any technical problems or other problems in doing activity at all.

**Structured Phone Interview with Teacher**

At the request of one of the teachers the interview was conducted over the phone. The other teacher was not available to interview. The teacher was concerned that 24 items in a portfolio was too much and was leading to her students spending too much time on the portfolio. This was exacerbated by students wanting to spend time redesigning earlier work that they felt they could now do better, when they were supposed to move to another task. Some students were spending too much time while others were concerned that they wouldn´t be able to measure up. She felt there needed to be a better balance of work and expectations.

The teacher said that many students had found it difficult to annotate work because they were not able to think at that level and therefore they relied on “teacher words” that they did not necessarily understand. Lower ability students could not cope with this at all. There were also difficulties with authenticating student work and ensuring equity. Some students had software at home and could do work at home and get others to help them while others could not. This limited the potential of the task as an external assessment. She felt that the range in quality of work related more to effort than skill of students and the help they received.

She would prefer a common task for external assessment or do a resume type of portfolio as the final production task. She would also like to see a practical exam like in Victoria.

**Results of Marking**

The portfolio output for 17 students were collected on a DVD, copied from the school’s server upon which the students saved their work. Each student had a folder containing folders and files that made up their portfolio. The formats of the files included FLV, HTML, DOC, PUB, MOV, JPG, GIF and PPT. Before marking could commence all the portfolios were checked, zipped and uploaded into the MAPS online.

The students’ work was marked by two external assessors using a standards-based rubric based on criteria associated with all four outcomes of the course, and by the teacher awarding a level for each of the four outcomes. For the standards-based marking method the assessors used the rubric provided in the course document based on the four outcomes, as described below. For the overall outcome and
each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the instructions in the table below on the criteria to assess. The portfolio was created either as a website, slideshow, Word document or Flash movie with own design for interface, including a range of work with explanations. This was collated towards the end of Semester One and would be further updated towards the end of Semester Two.

<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Technology Process: Students apply a technology process when creating or</td>
<td>Investigations – Develops proposals and shows consideration of alternatives.</td>
</tr>
<tr>
<td>modifying products using information communication technologies.</td>
<td>Planning and Communication – Clarity and detail of planning documentation, use of</td>
</tr>
<tr>
<td></td>
<td>diagrams, graphics and text.</td>
</tr>
<tr>
<td>1.1 investigate ideas, and evaluate alternative designs and proposals</td>
<td>Production, Evaluation and Management – Details of evaluation and production plans.</td>
</tr>
<tr>
<td>1.2 devise communicate and evaluate proposals</td>
<td></td>
</tr>
<tr>
<td>and design plans in appropriate forms</td>
<td></td>
</tr>
<tr>
<td>1.3 implement and evaluate production processes</td>
<td></td>
</tr>
<tr>
<td>and strategies to manage resources efficiently</td>
<td></td>
</tr>
<tr>
<td>2 - Understanding Information and Communication Technologies: Students</td>
<td>ICT concepts – Conceptual understanding of Digital Graphics (e.g. hardware, graphics</td>
</tr>
<tr>
<td>understand the nature and use of computer hardware and software to achieve</td>
<td>software) selection of ICT and techniques, and use of terminology.</td>
</tr>
<tr>
<td>information solutions.</td>
<td>Selecting &amp; Using ICT – Understanding of standards and conventions for information</td>
</tr>
<tr>
<td></td>
<td>solutions in Digital Graphics.</td>
</tr>
<tr>
<td>2.1 understand the ICT-related concepts, standards and terminology required</td>
<td>ICT Systems – Understanding of related components of ICT systems for digital graphics</td>
</tr>
<tr>
<td>to select and use appropriate computer software and hardware to achieve</td>
<td>(e.g. processor speed, file size, compression, resolution).</td>
</tr>
<tr>
<td>information solutions</td>
<td></td>
</tr>
<tr>
<td>2.2 understand the standards and conventions when selecting and using ICT</td>
<td></td>
</tr>
<tr>
<td>systems to achieve information solutions.</td>
<td></td>
</tr>
<tr>
<td>2.3 understand the management, processes, functions, types and relationships</td>
<td></td>
</tr>
<tr>
<td>between the components of ICT systems required to achieve information</td>
<td></td>
</tr>
<tr>
<td>solutions.</td>
<td></td>
</tr>
<tr>
<td>3 - Quality of Information Solutions: Students explore alternatives and use</td>
<td>Applying conventions – Applies standards and conventions for information solutions in</td>
</tr>
<tr>
<td>skills, techniques, processes, standards and conventions to achieve</td>
<td>Digital Graphics - spacing, colour, size, position, fonts, consistency, variations,</td>
</tr>
<tr>
<td>information solutions.</td>
<td>....</td>
</tr>
<tr>
<td>3.1 Apply accepted standards and conventions to create or modify various</td>
<td>Skills &amp; techniques – Applies efficient skills, techniques and processes in creating</td>
</tr>
<tr>
<td>forms of information solutions.</td>
<td>digital graphics (e.g. layers, tools, file format, editing)</td>
</tr>
<tr>
<td>3.2 Apply efficient skills, techniques and processes in the use of ICT</td>
<td>Enterprising – Creative and innovative designs or approaches to designs.</td>
</tr>
<tr>
<td>systems to develop information solutions.</td>
<td></td>
</tr>
<tr>
<td>3.3 Apply enterprising capabilities while exploring alternatives and working</td>
<td></td>
</tr>
<tr>
<td>to achieve information solutions.</td>
<td></td>
</tr>
<tr>
<td>4 - Information and Communication Technologies in Society: Students understand</td>
<td>Beliefs &amp; Values – Understanding of social impacts on digital graphics (e.g. perfect</td>
</tr>
<tr>
<td>how cultural beliefs, values, abilities and ethical positions are interconnected</td>
<td>images, ethics of editing, privacy).</td>
</tr>
<tr>
<td>in the development and use of information and communication technologies</td>
<td>Benefits &amp; Consequences – Understanding benefits and consequences of digital graphics</td>
</tr>
<tr>
<td></td>
<td>related to social impacts (e.g. access to images and tools to create images, easy to</td>
</tr>
<tr>
<td></td>
<td>misrepresent)</td>
</tr>
<tr>
<td>4.1 understand the cultural beliefs, values, abilities and ethical position</td>
<td>Impact – Understand consequences of digital graphics on individuals, groups and</td>
</tr>
<tr>
<td>that can impact on the use of ICT</td>
<td>environments (e.g. creates expectation of perfection, efficient communication)</td>
</tr>
<tr>
<td>4.2 understand the benefits and consequences of ICT use in different contexts</td>
<td></td>
</tr>
<tr>
<td>and how this relates to beliefs, abilities, values and ethical positions</td>
<td></td>
</tr>
<tr>
<td>4.3 understand the consequences of technological developments on social,</td>
<td></td>
</tr>
<tr>
<td>personal, cultural, physical, economic and ethical structures and environments</td>
<td></td>
</tr>
</tbody>
</table>
The two external assessors marked the student work on the aspects of the four outcomes while the teachers marked also using the four outcomes. The assessor judgements on the outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The numeric scores for each assessor were summed across the three outcomes to produce a score with a maximum of 100 for each assessor. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Consensus (C)</th>
<th>Teacher (%)</th>
<th>Rank A1</th>
<th>A2</th>
<th>C</th>
<th>Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA101</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>53</td>
<td>7.5</td>
<td>8</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 4.11: Tabulation of results from marking.*

A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

*Figure 4.12: Correlations between results from marking PA student work.*

The two external markers used the level statements for the outcomes to mark the portfolios. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.84 (p<0.001) on the ranking of students. This increased with some discussion of the results for students 106, 109, 111 and 112. On these students the main disagreement was on the level awarded for
Outcome 1. In some cases there was a difference in marks awarded because any judgement below L4 was awarded a 0 mark compared with a 12 for a low L4.

The teachers provided first semester marks based on an assessment on each of the four outcomes that was converted to a mark out of 25 and summed to provide a mark out of 100 for each student. The portfolio was not marked separately but all of the tasks collated within the portfolio were included in the semester mark.

Comparison with external marker results indicated little agreement apart from about four students. When compared with the two external assessors the inter-rater reliability correlations on the ranking of students were 0.21 and -0.01 respectively, with neither significant at the 0.05 level of significance. This was not surprising as the teacher mark was based on all submitted work in the semester, much of which most students had not adequately included within the portfolio.

Conclusions
Overall it was concluded that this digital portfolio was implemented with significant difficulty due to the lack of time available and the quantity of work students were required to collate. Further, for some this was a new experience and the choice of software to collate the portfolio consumed time for some students. There were no real technical difficulties in implementation.

Capturing the performance of the students was relatively difficult because, while each student’s work files were provided in folders on a single DVD with a folder for each student, the variety of software used and variety of states of completion made the material difficult to work with. After checking the technical operation of the portfolio, sometimes maintenance being required, each student’s folder was zipped as folder and uploaded into the MAPS portfolio system. It could then be downloaded for viewing by each marker. Where the portfolio was a website or Flash presentation file it could be viewed through the browser, where it was a slideshow or document it had to be viewed through the appropriate application.

The limited annotation provided by all students and the limited process documentation provided by most students made it difficult for students to demonstrate high level achievement and difficult for assessors to recognise that achievement. Where design documents were provided it was generally easy to assess Outcome 1 and often this provided some evidence for Outcomes 2, 3 & 4 but this usually needed additional evidence from examples of products or annotations. It was difficult to include the responses to test questions.

In terms of the marking of the student work the following conclusions could be drawn.

1. There were clear differences in results from the external markers and the teachers due to the teachers having access to all the student’s work. For most students large amounts of work were inadequately represented in the portfolio.

2. There was a high level of agreement between the two external assessors using the standards framework.

3. Only three students were judged to have demonstrated any work above Level 4 and nine students were judged by at least one assessor to not have demonstrated work at even Level 4 for all the outcomes. While it was likely that all the students judged not to have demonstrated Level 4 would have done so with time to properly collate the portfolios the three students judged to have demonstrated Level 5 work were not those indicated by the teachers.

4. To some extent it could be considered that the results from the external assessors related more closely to student organisation capability in collating a portfolio rather than their capability associated with all of the outcomes of the course.
MA: Video Production Exam

The MA case study involved one class of Year 12 students at one school completing the Industry Information Technology subject of study in a Digital Video context. There were 9 students in the class.

The aim was for students to research, devise, produce and evaluate a 30 second video advertisement against smoking. This was to be done individually, as much as possible under exam conditions, incorporating a full technology process over a two-week period in June using eight 45 minute sessions of class time.

Development of Assessment Instrument

The teacher developed the exam paper from scratch after discussion with the researcher and being provided with an example. The first ideas for the exam paper were emailed in January 2007 and were based on the requirements for a test in the Industry Information Technology subject that connected with outcomes related to knowledge of an industry, not practical production. In the beginning the researcher believed this mismatch led to a number of issues in the end such as the large amount of time spent on research. At that stage the researcher did not recognise this mismatch but did suggest completing all components of the exam on computer and reducing the output to three files (video, research and planning), however, later it was suggested to make this five files. By early May there had been little change to the exam paper. By the beginning of June the paper had progressed to its final form and now included some website references.

The researcher made a number of suggestions including:

- Consistent naming of the files to save (implemented)
- All components to be presented on computer (implemented)
- Some changes in wording (largely implemented)

The teacher and researcher discussed what materials students should have access to with the result that they were permitted access to the Internet and were given website references. In addition they were given access to a video camera for two sessions each.

Originally the video was to be 90 seconds but it was decided to reduce the time to reduce the file sizes and force students to be more selective.

At the beginning the following structure was considered.

You will complete the following tasks during the time provided in class.

1) PROPOSAL: A proposal for the development of the video including a definition of the industry and the area of the challenge.
2) DESIGN DOCUMENT: A design document for the video that may include drawings, flowcharts and schematics.
3) PROCESSES PLAN: A plan of project activities, sequencing and logistics.
4) PROTOTYPE: A prototype of the video indicating all significant components.
5) REFLECTION: Responses to questions that reflect on the processes and outcomes. This will be completed as a separate electronic test.

The output from the first FOUR tasks will be collated into an electronic folder on a DVD and must be accessed through <NAME OF SOFTWARE>.

Implementation of Exam

Students were given an electronic copy of the exam paper and the teacher invigilated. The instructions on the exam paper were as follows:

You will be given two weeks to create a max 30 seconds advertisement for an Anti-smoking campaign. It is an individual task and requires you to plan, film, edit and produce a final copy advert
to be assessed by a marketing executive. It is open book and you can use any material that you have used/found in class. You will need to create the following items as part of your test:

- Planning Storyboard considering all good design principles of Video Creation
- Characters, shots, lighting, props
- Edited footage
- Story of Creation
- Research completed on Smoking facts and figures
- Research into film making industry

The Test is a practical test and will use IT and paper to complete it. You will have access to a video camera for two allocated periods of time (45 minutes each time). All your work should be stored in electronic form in a single folder or on a DVD. Paper documents should be scanned and placed in the folder. The folder will be named using your student ID.

1. (1) RESEARCH (Save as a document called **research**)  
Discuss the results of your research into:
- Smoking facts and figures.
- The film making industry.
Reference any documents or websites you used.

2. (2) PLANNING (Save as a document called **storyboard**)  
Develop a design document for the video that will include a STORYBOARD that includes CHARACTERS, SHOTS, LIGHTING and PROPS.
- The document should include adequate information that would allow another skilled person to complete the project.

3. (3) PROTOTYPE VIDEO: (Save as a file called **video**)  
You will produce and provide a prototype of the video. This will be provided in one or more electronic files on the DVD.
- You will apply at least two pieces of computer software in the project.
- Include technical and user manuals that are accessed through a web browser.

4. (4) STORY OF CREATION: (Save as a document called **mystery**)  
Write a brief report on the development of your video.
- An outline of the human need or opportunity that you addressed.
- The main objectives of the project.
- Briefly document all resources applied to the project: hardware, software, and materials.
- Include a list of skills and understanding in applying appropriate pieces of hardware and software.
- Briefly describe how you created the video. Give examples of early attempts which were improved with explanation of the improvements.
- Explain how well you think you did in achieving your objective.
- Explain what you would change if you did it again.

5. (5) REFLECTION: (Save as a document called **reflection**)  
You will respond to a series of reflective questions on the processes and output of your project. The content would include questions concerning:
- the technology process employed.
- the key-points of the design.
- design ideas, tools and techniques.
- justification of ideas and resources.
- comparisons with processes and techniques used in the industry
- social issues related to the solution and the processes of the project
Results
A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their exam.

Observations of the class
The class of 10 students were observed for one session in each of the two weeks of the exam. The class was in a ‘fish bowl’ room so there were some distractions from other classes.

Session 3: 8 students present for 45 minutes.
They had spent the previous 2 sessions on research. Teacher was concerned that they were taking a long time so after consultation with researcher decided to suggest to students that they start planning and stop researching. In addition they were to be permitted to use existing video and graphics that they could download off the web. At the beginning of the session 2 were working on video using Premiere (mainly downloaded video and title screens), 2 planning in Word using ideas off the web and a paper-based storyboard sheet, 4 were copy and pasting from research using Google.
All were quiet and focussed (teacher said they were not yesterday). Some conversations earlier between students regarding content and design ideas. Some students looked at each other’s ideas. A student asked the teacher for a sound effect file so teacher put some on a server. There was only 1 female, she was using the paper storyboard all session and ideas off web research. One male was selecting sound effects and another asked teacher how to edit a graphic. No evidence of referencing materials.
After 30 minutes, 4 were working in Premiere, 1 Photoshop, 1 storyboards, 1 Word planning, and 1 Word research. Students had their own folders but these were not secure. Only 1 student was ready to film. Teacher indicated that they could use students from other classes in their video.

Session 5: Only 7 students present for 45 minutes. Teacher will be away for rest of week.
No students wanted to use video camera at all they would just use Google downloaded video and stills in Premiere. They were permitted to have extra ‘study’ periods working on it in the laboratory (they were asked to record how much extra time they took).
All focussed but some conversations and viewing of the work of others. Some use of Photoshop to edit downloaded photos. Only 1 student worked off a storyboard. 2 students had background music playing. Two still finding media using Google.
After 20 mins the Internet access was interrupted. There was a lot of work on title screens. All students appeared comfortable using Premiere. One student showed another how to create an effect in Premiere. Another student wanted a cigarette to film so rolled up paper (teacher helped). There were three more sessions to go (didn’t look like they will finish - 45 minute sessions appear to be too short).

Survey of Students
There were 8 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.
Table 4.5
Descriptions and descriptive statistics for the scales based on items from the student questionnaire (N=8).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>8</td>
<td>2.8</td>
<td>3.9</td>
<td>3.4</td>
<td>0.40</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>8</td>
<td>1.6</td>
<td>2.6</td>
<td>2.1</td>
<td>0.32</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>8</td>
<td>2.0</td>
<td>2.6</td>
<td>2.3</td>
<td>0.18</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confidence</td>
<td>8</td>
<td>1.8</td>
<td>2.7</td>
<td>2.4</td>
<td>0.28</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>8</td>
<td>2.6</td>
<td>3.8</td>
<td>3.2</td>
<td>0.45</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>8</td>
<td>77</td>
<td>132</td>
<td>110</td>
<td>16.36</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

Figure 4.13: Distribution of scores for the Skills and SCuse scales on the student questionnaire.

From the table above it is clear that the students didn’t tend to have done the nominated activities in Q10 and were not confident they could do them except by typing an assignment for school that they all had done. Most liked using computers at school and home but only sometimes liked finding out things for themselves. Seven of the eight indicated great confidence in using computers. Seven had broadband Internet access at home and five used computers on a daily basis.

Almost all Agreed or Strongly Agreed with Q2 items concerning value of eAssessment (only two disagreed on any of these items). Most had little experience with tests on computers but felt that they could quickly get used to it.

Students indicated that the assessment was relatively easy to complete, that they typically had not done the range of applications indicated, that they were fairly confident in using ICT, and had a high level of ICT skills across the range of types of applications. On average they indicated using ICT for nearly two hours per day at school with the lowest 77 minutes and the highest 132 minutes.

There were two open-ended questions on the student questionnaire. Item 3 asked students to list the two best things about the assessment and Item 4 asked them to list the two worst things about the assessment. Generally the students considered that using computers allowed access to more information and tools and was easy to manage. The main worst things were limitations of time and facilities and one student did not like using computers.
Students Forum

The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.

The students indicated that they thought doing the exam was “good”, as it allowed them to do express their own creative thing, tackle an issue in the media and have their own interpretation for two weeks. They explained that they had to include the use of Adobe Premiere but didn’t actually have to include any filming process, because they were allowed to use clips from the Internet. They believed that all students in the class would have liked doing the exam instead of the paper exams they had always done in the past.

They believed that they had been able to produce a high standard of work because they had all the tools they wanted. However, they felt that they needed a bit more time, particularly to do the “paperwork” as well as the research on the topic. These two areas severely reduced the time in the two weeks available to edit their video. They indicated that their time was not really structured, and they were given the responsibility of organising their time. They thought that the topic of smoking was a good topic because there was a lot of “stuff” available for them to save onto the network.

They were concerned that technical problems could occur leading to the loss of work. However, they could only cite one example of another student who they “thought” had a couple problems with his project that resulted in lost information. They felt that the danger of losing things was a problem with doing an examination digitally. They were aware that computers may freeze and that this was a hardware problem and not a software problem. They were concerned that they only had 45 minutes per session giving 12 hours over two weeks and they could not work at home because they did not have access to the software or their files at home. They could not identify any other types of problems in completing the exam.

Structured Email ‘Interview’ with Teacher

The teacher believed that the task worked well and the students were clear in what was required. However, he thought that there was too much to do in the time available but that the “level of detail will need to be maintained in an external exam”. He felt that the students could not do their best and have time to film with the amount of work that was to be done, although the sub tasks were necessary and this affected the better students more. He was confident that the students preferred an electronic based exam over a written one, particularly the boys and that they tended to do the practical sections well and not the written response sections. He indicated that for the time he was reasonably happy with most of the work, however, time management was an issue, due to students placing more time on research and the practical and rushed the last couple of documents, with evaluations being a little elementary. He was surprised at how well they worked individually but was disappointed with the quality of some students work particularly the written sections.

While there were no real technical problems with implementing the activity the internet was relied upon quite a bit so if that had gone down, students would have been at a loss as for what to do. There was a problem with the teacher being absent for a few days that he felt may reduce the validity of the test. He believed that while the type of exam had potential it would not be possible as a 2 week external exam but that a 2-3 hour practical test would be possible with limits to the task and not as in-depth.

Results of Marking

The exam output for 9 students were collected on a DVD, copied from the school’s server upon which the students saved their work. Each student had a folder containing files generated from their exam. The formats of the files included AVI, HTML, DOC, JPG, and GIF. Before marking could commence all the files were checked, the required Word documents were combined into one PDF file and this, along with the AVI movie file was uploaded into the MAPS online portfolio system.
The students’ work was marked by two external assessors using a standards-based rubric based on criteria associated with Outcomes 1 and 3 of the course, and by the teacher awarding a level for Outcomes 1, 4, 6, 9 and 10 of the Industry Information Technology subject. For the standards-based marking method the assessors used the rubric provided in the course document based on the two outcomes, as described below. For the overall outcome and each “Aspect” of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the instructions in the table below on the criteria to assess. The extended production exam output was presented to the assessors in two digital files: (1) an AVI file titled Video, a prototype of anti-smoking advertisement video produced in Premiere; and (2) a PDF file comprising the following four Word documents.

**Research Doc:** Research smoking facts & film making  
**Storyboard Doc:** Storyboard for a video  
**Mystery Doc:** Report on development of video  
**Reflection Doc:** Reflection on whole project

The criteria used by the assessors are described in Table 4.6.

**Table 4.6**  
Description of criteria used in assessing work produced by MA students.  

<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
</table>
| 1 - Technology Process: Students apply a technology process when creating or modifying products using information communication technologies.  
1.1 investigate ideas, and evaluate alternative designs and proposals | Research Doc – quality of investigation to find information and smoking and video advertising industry.  
Mystery Doc - Explanation of how investigation influenced design and production (e.g. early attempts & changes).  

| 1.2 devise communicate and evaluate proposals and design plans in appropriate forms | Storyboard Doc – quality of communication of design - characters, shots, lighting, props.  
Mystery Doc - Explanation human need and project objectives in relation to design.  

| 1.3 implement and evaluate production processes and strategies to manage resources efficiently | Mystery Doc – quality of implementation of production processes and evaluation of investigation, design and production processes.  
Reflection Doc - Evaluation of processes, design and production. |
|----------------|---------------------|
| 3 - Quality of Information Solutions: Students explore alternatives and use skills, techniques, processes, standards and conventions to achieve information solutions.  
3.1 Apply accepted standards and conventions to create or modify various forms of information solutions. | Video – Application of standards and conventions for digital video public adverts (grabs attention, conveys message, smooth transitions). Consideration of audience in design (e.g. youth, women).  
Reflection Doc - Explanation of design features, justification of ideas and resources.  

| 3.2 Apply efficient skills, techniques and processes in the use of ICT systems to develop information solutions. | Video –Application of digital video editing skills such as combining, transitioning, audio tracks, effects. Whether included own footage. Message connected with effects.  
Mystery Doc - Explanation of skills and techniques used.  

| 3.3 Apply enterprising capabilities while exploring alternatives and working to achieve information solutions. | Video –Shows some initiative and individuality - e.g. includes own video, sitsls, effectiveness of message.  
Mystery Doc - Explanation of rationale for design features that are creative. |
The outcomes addressed by the teacher were outcomes 1, 4, 6, 9, and 10 of the Industry Information Technology subject and use the related descriptors for Satisfactory, High and Very High performance. These outcomes align best with Outcome 1 and 3 of the new Applied Information Technology course and therefore the external assessors used these two outcomes in their assessment of the work.

Outcome 1: Investigates and analyses the impacts of current and future information technologies in industry contexts.
- evaluates the social, ethical, environmental and personal consequences of investment
- explores the impact of trends in development of information technology
- explores the impact of information technology on work practices in industry

Outcome 4: Considers the design principles of information technologies used in a variety of industry contexts.
- recognises the need for design principles
- describes design principles when designing solutions in an industrial context
- describes design issues relating to the safe and efficient use of the technology
- describes a range of techniques used to improve the human-computer interface.

Outcome 6: Describes and evaluates information technologies (products, systems, processes, services and environments) used in industry contexts.
- identifies the elements of information technologies used in industry contexts
- identifies the components of information technologies
- describes the functions of the components of information technologies used in industry contexts
- describes the relationship between the components of information technologies
- compares alternative information technologies used for similar purposes in industry contexts.

Outcome 9: Describes the effect of information technologies on personnel, work roles, training and employment opportunities within specific industries.
- discusses work roles that involve information technology in industry
- carries out research on how information technology affects employment in specific industries
- describes the changing employment opportunities resulting from the application of information technology industry

Outcome 10: Demonstrates an understanding of the application of computer communications systems to workplaces.
- describes the functions of hardware components used for computer communication
- describes aspects of computer communication
- demonstrates understanding of computer communication technologies.

The two external assessors marked the student work on the aspects of the two outcomes (1 & 3) while the teacher marked using the CAF Course Outcomes 1, 4, 6, 9 and 10 and the grades ND, S, H and V.

The assessor judgements on the two outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The two assessors met after they had marked the work and came to a consensus mark for each outcome for each student. The numeric scores for each assessor were summed across the two outcomes to produce a score with a maximum of 50 for each assessor. For each assessor, for their consensus mark and for the teacher’s mark the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Consensus (C)</th>
<th>Teacher (%)</th>
<th>Rank A1</th>
<th>Rank A2</th>
<th>Rank C</th>
<th>Rank Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA101 etc.</td>
<td>17</td>
<td>14</td>
<td>16</td>
<td>20</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

*Figure 4.14: Tabulation of results from marking.*
A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

![Correlation Table]

*Figure 4.15: Correlations between results from marking MA student work.*

Two external markers used the level statements for the three outcomes to mark the exam outputs. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.70 (p=0.05) on the ranking of students. They disagreed substantially on only two (MA102 and MA104) of the 9 students (Level 4 cf below Level 4 that attracted a 0 mark) due to slight differences in assumptions made about students’ work when limited documentation was provided (e.g. output from investigation was largely copied from websites). However, there were some differences on individual outcomes between Level 4 and 5, in particular for MA109 where the emotive content of the video influenced the assessors in opposite ways. Also there was a difference in opinion on whether it was important that students include original video footage or edit downloaded clips and stills as was the case for all students. It is therefore important that the relationship between the requirements of the tasks and interpretation of the outcome statements be made more explicit and that students be afforded increased scaffolding to ensure they indicate how their product was developed and the sources of material used.

The teacher used the CAF Course Outcomes 1, 4, 6, 9 and 10 and the grades ND, S, H and V and also intermediary grades of ND/S, S/H and H/V. These grades were converted to a number (ND=0, ND/S=1, S=2.5, S/H=3, H=4, H/V=4.5, V=5) and then the five values summed for each student to give a single score that could vary between 0 and 25. This was then multiplied by four to give a percentage.

Comparison with external marker results indicated good agreement apart from the students MA102 and MA104 with Assessor 1. When compared with the two assessors the inter-rater reliability
correlations on the ranking of students were 0.72 and 0.84 respectively, with both significant at least at the 0.05 level of significance.

**Conclusions**

Overall it was concluded that this computer-supported production exam was implemented with no significant technical difficulties but with most students not completing all the requirements of the exam in the two weeks of time permitted. It is likely that the students spent too much time searching for information and audiovisual content and also were less interested in completing ‘written’ components of the exam such as planning and evaluation documents. The students responded well to the style of exam and generally preferred it to a theory paper-based exam although they mainly liked creating the video and finding the content.

Capturing the performance of the students was relatively straightforward by copying their work folders onto a DVD, converting all documents into one PDF file for each student and uploading these and the video file into the MAPS portfolio system. For one student some hand drawn storyboard pages had to be scanned and included in the PDF file. The structure of the assessment task made it easy to collate and mark a student’s work. The lack of time spent by students on the planning, evaluating and reporting documentation made it difficult for students to demonstrate high level achievement and difficult for assessors to recognise that achievement. No students included any original video material and most students provided very little connection between the research, planning and production stages. It is likely that they needed additional scaffolding to encourage this to occur, for example responding to questions or filling in boxes using points or lists rather than paragraphs.

In terms of the marking of the student work the following conclusions could be drawn.

1. There were clear differences in results based on marking by outcomes using the standards framework compared with the outcome descriptors for the IIT subject. This was to be expected given the differences in what the outcomes are designed to measure. However, despite this the final ranking of the students was almost identical.

2. There was a high level of agreement between the two external assessors using the standards framework with only two students being ranked significantly differently. At a consensus meeting these differences were overcome.

3. Only three students were judged to have demonstrated any work above Level 4 with only one judged to be so by both Assessors. Four students were judged by at least one assessor to not have demonstrated work at even Level 4 for Outcome 1, mainly due to inadequate documentation.
CA: Web-Page Production Exam

The CA case study involved one class of Year 11 students at one school completing the Applied Information Technology course Unit 2A in an ICT in Sports context. There were 10 students in the class who completed a 3-hour exam on one day in June.

The aim was for students to create a web page and respond to questions around a given scenario, done under exam conditions. The focus of activity was on aspects of the technology process, information design, and ICT knowledge and skills. The assessment task centred on a major scenario (Tennis Club) with a challenge (Promote and organise enrolment in a one-day tournament) that enabled students to develop an appropriate design and prototype solution. The scenario was part of an on-going project the students had been involved with during the term with the challenge to add a component to a website. The students also were required to respond to a series of design-related questions and evaluate the design of their web page.

Students would demonstrate knowledge and skills in web-page production and information design and storage through completing a web page and responding to questions about the design and implementation of the information solution, and evaluating their prototype.

Development of Assessment Instrument

The teacher developed the exam paper from scratch based on work being completed by students during the year to create a web site for a tennis club. The first ideas for the exam paper were emailed in February 2007 as less than one page in three sections: design and produce a web page; describe hardware and software; and consider social implications. By the end of March the paper had progressed to almost the final form.

The researcher made a number of suggestions including:

- Consistent naming of the files to save (implemented)
- Using the terms ‘Design Questions’ and ‘Evaluation Questions’ (implemented)
- Some changes in wording (largely not implemented)
- Increase mark allocation to prototype (implemented)
- Investigation prior to the exam (not implemented)

The teacher and researcher discussed what materials students should have access to – the teacher decided to use secure logons that prevented access to network facilities but still allowed access to software application help files. This did not allow students to access their own work on the web site to use graphics and check for consistency.

Initially the exam was set for 2 hours but the teacher decided to extend this to the maximum 3 hours.

Originally the researcher had suggested using a scaffolded approach to design including: Proposal, Design Document, Prototype, and Evaluation/Reflection. The teacher probably felt that the students would find this too difficult to complete within the time allowed.

Implementation of Exam

Students were given a paper copy of the exam paper (3 pages) and a double-sided A3 sheet with level statements for Outcomes 2 and 4. There was 10 minutes reading time prior to the 3 hours of work. Each student was provided with a USB Flash drive.
The instructions on the exam paper were:

Create a folder in the USB drive in which to place all your exam files. The directory should include your student key followed by AIT2A (e.g. jonesqAIT2A)

Read the narrative on the next page

Answer the Design Questions in Design_questions.doc

Develop the prototype for the data entry page in any web page editor

(Create a data entry form to allow participants to enter their details. The overall look of the form should match your website. You DO NOT have to include your narrator)

Complete the evaluation

Special logins had been created for the exam so that there was no Internet access (e.g. CA101, Password1). The teacher and researcher invigilated the exam.

Results

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class

The class of 10 students completed the exam on one day from 9am to 12pm with 5 minutes reading time before this (could also login during this time).

Special logins had been created for the exam so that there was no Internet access (e.g. CA101, Password1). Students were given a paper copy of the exam paper (3 pages) and a double-sided A3 sheet with level statements for Outcomes 2 and 4.

Two students began on the web page using FrontPage while the others started answering the design questions using Word. None saved to start with and after 12 minutes still none had saved so they were reminded to do so. After 5 minutes one student stopped using FrontPage and went to answering the questions. After 20 minutes 4 were working using FrontPage with 2 stopping at 45 minutes and 1 starting. After 50 mins 1 had started using Paint for a graphic and 6 were using FrontPage. One student wanted to access his website to match the ‘look and feel’ but was blocked. One student asked how to create a home page. After 60 mins 5 students doing web page and one using Flash for graphics. At 105 mins 1 student started editing in HTML using FrontPage and continued to do so for most of the rest of the time, towards the end one other student started editing using HTML (no other students did at any time). One student succeeded in accessing the network and finding an online textbook.

Two students appeared to finish 30 mins early but kept working. After a while one started playing a game and progressively students finished and started playing games until at the end only 2 students were still working on the exam. None of the students were observed referring to the outcomes sheets.

Survey of Students

There were 8 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below with scores calculated so that high scores are more positive.
Students indicated a high level of ICT skills across the range of applications. On average they indicated using ICT for about one hour per day at school with the lowest 42 minutes and the highest 96 minutes. They didn’t tend to have done the nominated activities in Q10 but were confident they could do them. All had broadband Internet access at home and used computers on a daily basis. Almost all Agreed or Strongly Agreed with Q2 items concerning value of eAssessment (only 2 disagreed on any of these items).

There were two open-ended questions on the student questionnaire. Item 3 asked students to list the two best things about the assessment and Item 4 asked them to list the two worst things about the assessment. Generally students considered that using computers made it easier, was a better environment, allowed access to more information and tools and was easy to manage. Main worst things were sore eyes and taking time if you are not experienced with computers.

Students Forum
The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.
The students indicated that they liked the practical part of the examination. They liked to do the web page design but they said the reaction of other students was that they often didn’t have enough background knowledge to answer some of the questions that were asked in the second part. They relied on their own general knowledge and not what they had been taught in class. However, they felt that they had been able to do their best quality work and they didn’t have any problems. They thought the practical IT was great for the examination and previously they had just done the paper exams.

They indicated that some things could be changed in the future. They felt they were limited in the web page design. They couldn’t use images from any libraries or the Internet and the tool they used, “Frontpage”, they didn’t like very much. They thought it was primitive and didn’t give them enough flexibility and would like access the things like image tools and more freedom in the design beyond the limited constraint of the graphic tools that they had in Frontpage.

They could not recall any technical problems in doing the activity at all, and because there were only a few students in the class they could sit one computer apart and thus not look at the work of other students. When asked about other suggestions to develop this type of assessment, they reinforced only that they preferred the practical assessment. They like doing practical things as that’s why they chose the subject.

**Structured Email ‘Interview’ with Teacher**

The teacher believed that the task suited both his needs as a genuine assessment tool and the administration’s requirement for a Semester Exam. He felt that the delivery had worked well and there were no significant problems. He was confident that the students preferred an electronic based exam over a written one as they felt this gave them a better chance of demonstrating what they knew and what they could do. As such, this type of exam has considerable potential. He believed that the only major improvement that could be considered was to gather the output files centrally on a server, this would also provide backup to the USB drives. However, he felt that to be more broadly implemented would require “more fail safe mechanisms”. He was pleased with the quality of the students’ work and was pleased that the students had responded so well to this style of exam.

**Results of Marking**

The exam output for 10 students were collected on USB Flash Drives, upon which the students saved their work. Each student had a folder containing folders and files with the formats of the files included HTML DOC, JPG, and GIF. Before marking could commence all the files were checked, the website folder zipped and uploaded along with the Word document containing answers to written questions, into the MAPS online portfolio system.

The students’ work was marked by two external assessors using a standards-based rubric based on criteria associated with three of the outcomes of the course, and by the teacher awarding a level for each outcome. For the standards-based marking method the assessors used the rubric provided in the course document based on the four outcomes, as described below. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the instructions in the table below on the criteria to assess. There were two files associated with the 3-hour exam: (1) a web-page that could be added to a web-site developed during the term; and (2) a Word document comprising responses to SIX design questions and THREE evaluation questions.
<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors' Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 - Understanding Information and Communication Technologies: Students understand the nature and use of computer hardware and software to achieve information solutions.</td>
<td><strong>2.1 understand the ICT-related concepts, standards and terminology required to select and use appropriate computer software and hardware to achieve information solutions</strong>&lt;br&gt;Design Question 1 – Level of understanding from description of computer configuration to host website (hardware, software, network devices, back up). Justification of configuration selected, particularly in terms of concepts, standards, terminology and needs of users and clients.</td>
</tr>
<tr>
<td><strong>2.2 understand the standards and conventions when selecting and using ICT systems to achieve information solutions</strong>&lt;br&gt;Design Question 2 – Expects use of standards and conventions for design of web page form and relates this to likely users (e.g. order &amp; position of fields, text and graphics). Explains how fits within the look-and-feel of their website and how it is differentiated to attract members and organise the tournament.</td>
<td></td>
</tr>
<tr>
<td><strong>2.3 understand the management, processes, functions, types and relationships between the components of ICT systems required to achieve information solutions.</strong>&lt;br&gt;Design Question 3 – Indication of understanding of techniques and procedures for getting data from web pages, storing and retrieving, and relationship to types of software.</td>
<td></td>
</tr>
<tr>
<td>3 - Quality of Information Solutions: Students explore alternatives and use skills, techniques, processes, standards and conventions to achieve information solutions.</td>
<td><strong>3.1 Apply accepted standards and conventions to create or modify various forms of information solutions.</strong>&lt;br&gt;Evaluation Question 1 – Layout design considered the audience and accepted standards and conventions for online forms for the type of data needed. &lt;br&gt;Evaluation Question 2 – Fields considered purpose of form (identify new members, email them, organise tournament), knowledge of users and type of data. &lt;br&gt;Prototype – Online form collects adequate data to identify new members, email them and organise tournament for a range of people. Layout of form suits novice users and tennis players. Layout is consistent with standards and conventions for online forms.</td>
</tr>
<tr>
<td><strong>3.2 Apply efficient skills, techniques and processes in the use of ICT systems to develop information solutions.</strong>&lt;br&gt;Prototype – Use of features available in Front Page to present form (e.g. fonts, entry types) and layout of page including text, graphics, and links (e.g. positioning).</td>
<td></td>
</tr>
<tr>
<td><strong>3.3 Apply enterprising capabilities while exploring alternatives and working to achieve information solutions.</strong>&lt;br&gt;Design Question 3 – Level of sophistication and creativity in method of data storage and retrieval and software for the task. &lt;br&gt;Evaluation Question 3 – Extent of creative design to cater for variety of users and purpose of info product.</td>
<td></td>
</tr>
<tr>
<td>4 - Information and Communication Technologies in Society: Students understand how cultural beliefs, values, abilities and ethical positions are interconnected in the development and use of information and communication technologies</td>
<td><strong>4.1 understand the cultural beliefs, values, abilities and ethical position that can impact on the use of ICT</strong>&lt;br&gt;Design Question 4 – Understanding of legal constraints in gathering and storing the data provided by users (e.g. explaining use of data, why it is needed, how it will be stored and destroyed, ensuring users know they are giving consent and that they are old enough to do so). Understanding demonstrated in suggested actions to secure data (e.g. protected server, authorised personnel access) and in consideration of other ethical issues (e.g. destroy data if they don’t become a member, don’t promise what can’t be delivered).</td>
</tr>
<tr>
<td><strong>4.2 understand the benefits and consequences of ICT use in different contexts and how this relates to beliefs, abilities, values and ethical positions</strong>&lt;br&gt;Design Question 5 – Explanation of benefits and consequences of online data collection show an understanding of relationship to context and the beliefs, abilities and values of users, clients and themselves (e.g. easy collection of data and publicity that clients will value, easy and non-threatening access for some users but not for others depending on their abilities and beliefs re online, impression created by web page match matches club).</td>
<td></td>
</tr>
<tr>
<td><strong>4.3 understand the consequences of technological developments on social, personal, cultural, physical, economic and ethical structures and environments</strong>&lt;br&gt;Design Question 6 – Explanation shows understanding of consequences of online presence for the tennis club and other clubs (e.g. savings of advertising and data collection, 24-7 presence to the world, probably outsource data handling so lose some control of data, need to maintain information – work pressure).</td>
<td></td>
</tr>
</tbody>
</table>

The assessment addressed all aspects of Outcomes 2, 3 and 4 of the Applied Information Technology course and used the levels of achievement. This was made explicit to the students in the exam paper. However, the teacher marked the exam using a traditional cumulative marks approach with marks indicated to students on the paper. Only 25% of the marks were for the web page prototype.
The two external assessors marked the student work on the aspects of the three outcomes (2, 3, & 4) while the teacher marked analytically to give a mark out of 100 (25 for the website and 75 for responses to the questions). The assessor judgements on the three outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The numeric scores for each assessor were summed across the three outcomes to produce a score with a maximum of 75 for each assessor. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1 Initial</th>
<th>Assessor A1 Final</th>
<th>Assessor A2 Initial</th>
<th>Assessor A2 Final</th>
<th>Teacher (%)</th>
<th>Semester (%)</th>
<th>M1</th>
<th>M2</th>
<th>T</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA101</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>60</td>
<td>70</td>
<td>6</td>
<td>5.5</td>
<td>9</td>
<td>etc.</td>
</tr>
</tbody>
</table>

*Figure 4.17: Tabulation of results from marking.*

A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

*Figure 4.18: Correlations between results from marking CA student work.*

Two external markers used the level statements for the three outcomes to mark the exam outputs. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.95 (p<0.0001) on the ranking of students. This increased with some discussion of the results for students 103, 104 and 105. Initially they disagreed substantially on only one (CA104) of the 10 students (Level 4 cf Level 5/6) due to a difference in the weighting placed on the written responses compared with inferences made from the prototype. Note that the difference in result for student CA107 is based on an award of N/A (0) compared with the lowest award of Level 4 (12) for one outcome.

The teacher used the numerical marks allocated to each component of the exam to provide a mark out of 100 for each student. The teacher also provided the semester total percentage for each student.
Comparison with external marker results indicated little agreement apart from the top three ranked students. In fact the bottom two ranked students by the assessors were ranked 5/6 by the teacher’s marking. When compared with the two assessors the inter-rater reliability correlations on the ranking of students were 0.45 and 0.47 respectively, with neither significant at the 0.05 level of significance.

Conclusions
Overall it was concluded that this computer-supported production exam was implemented with no significant technical difficulties with all students able to complete the requirements of the exam in the time permitted. It is likely that the students did not need a full 3 hours to complete the exam, possibly 2 hours would have been adequate. The students responded well to the style of exam and generally preferred it to a theory paper-based exam although they mainly liked creating the web-page, not answering the questions. Some felt that they had not been introduced to the content required to respond to some of the questions. Capturing the performance of the students was relatively straightforward using the USB drives, zipping the web folder for each student, and uploading this along with their two Word documents into the MAPS portfolio system.

The structure of the assessment task made it difficult for students to demonstrate high level achievement and difficult for assessors to recognise that achievement. This was particularly due to the extended response questions and the limited nature of the website development. The Design Questions tended to be general and abstract and not well enough connected with the main design task. If students were required to respond to the questions using points or a list rather than a paragraph format it would have been easier for the assessors and would have left students more time to work on the website.Verbose answers may limit opportunity to demonstrate higher order thinking. Perhaps impose a word limit on answers. The questions represented a very high proportion of the marks for the examination and therefore compelled students to spend a long period of time on them leading to very repetitive responses.

In terms of the marking of the student work the following conclusions could be drawn.

1) There were clear differences in results based on marking by outcomes using the standards framework compared with analytic numeric scoring with the difference most obvious for lower achieving students.

2) There was a high level of agreement between the two external assessors using the standards framework.

3) Only two students were judged to have demonstrated any work above Level 4 and four students were judged by at least one assessor to not have demonstrated work at even Level 4 for at least one outcome.
XA: Poster Graphics Task Exam

The XA case study involved two classes of Year 11 students at one school completing the Applied Information Technology course. There were 49 students in the classes but three students were not involved in the study because either they did not want to be involved or their parents did not give permission.

The aim was for each student to complete a poster using some graphic images that were provided and answer 10 multi-choice questions about digital graphics. The focus of activity was on applying technical digital graphics skills in producing the posters not on the technology process. No scenario or requirement to address a particular audience was provided. The teacher developed the assessment task, based on a previous test paper, to suit his own requirements and decided to mark the exam using the statements for three of the learning outcomes of the course and a cumulative marks approach with marks mapped back to levels of achievement. Ten marks were allocated to a set of 10 multiple-choice questions and 80 marks were allocated to the production of the digital poster. This was developed at least a month prior to the exam. The students had done two practical tests earlier in the year. During the same period students were to complete a paper-based theory exam for the course. The researcher visited the classes to observe them working on one of the practical tests.

Students were given an electronic copy of the exam paper (3 pages) through an online Learning Management System and permitted to use any software available on their desktop computer. There was 5 minutes reading time prior to the 50 minutes of work under exam conditions (no communication with other students). Each student saved work to his/her own server space. Students had full access to computer network and could choose to use any software they wished. They were required to use at least four of the graphics supplied in the exam paper (Word document). The teacher invigilated. The instructions on the exam paper were:

Using at least four or more of the eight (8) graphics below, plus your own photo - create an advertisement about any topic use any suitable software. Save as your initials_graphic to your EXAM FOLDER.

Results

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their examination.

Observations of the classes

A researcher observed the two classes completing the exam, on the same day, in consecutive time slots.

The first class had 24 students. Due to a school event the class was 25 minutes late starting so the students worked into recess time with the teacher offering muffins as an incentive. There was some unhappiness about this and some discussion with the teacher about unfairness but the teacher prevailed and the students started work quickly. One student’s computer was slow to log on so she was given an extra 5 minutes (seemed to be the technically best student). A few students were up to an additional 6 minutes late arriving. Students could easily view their neighbour’s screens but after a few minutes there were no obvious instances of this happening. All students quickly focused and generated ideas for the poster (most started with the poster and went back to the multiple choice questions). They used a variety of software but each only used one (50% Corel PhotoPaint, 20% Publisher, 20% Fireworks, Powerpoint, CorelDraw). A few used templates to start (e.g. Publisher) and two students got graphics using Google. Two students did not have a photo of themselves so had to do their best without. One student spent a lot of time editing his own photo, the rest did not. Only one student needed technical help with saving. Six students finished after 35 minutes and only four remained to the end.

The second class had 22 students and started the exam on time. Some students were concerned about the weighting compared with other assessments and felt 10-15% was too high. A few had trouble finding the exam paper and asked the teacher for help. All started on the multiple choice questions with five students spending a lot of time using Google to help answer the questions. Quite a few used
Word for the poster, 50% used Photo Paint, a few used Publisher, 2 used CorelDraw, 1 Powerpoint and 1 started on Fireworks, then switched to CorelDraw and then Powerpoint. Overall they seemed to be a lower skilled class. One student who worked in Word did edit one graphic in Photo Paint. Generally students focussed on the task though less so than the other class. Students started finishing 15 minutes early and then either played online games or worked on other schoolwork. Some spent a lot of time on single images, cropping and filtering. At least 5 finished 10 minutes early but 13 went to the end. Two boys sitting next to each other used the same ‘Google’ graphic.

**Survey of Students**

There were 44 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.

**Table 4.8**

<table>
<thead>
<tr>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>2.7</td>
<td>4.0</td>
<td>3.4</td>
<td>.34</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>1.6</td>
<td>3.0</td>
<td>2.4</td>
<td>.39</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>2.0</td>
<td>2.8</td>
<td>2.4</td>
<td>.19</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confidence</td>
<td>1.7</td>
<td>3.0</td>
<td>2.5</td>
<td>.24</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>2.3</td>
<td>4.0</td>
<td>3.4</td>
<td>.44</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>16</td>
<td>468</td>
<td>110</td>
<td>81.26</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

**Figure 4.19: Distribution of scores for the Skills and SCuse scales on the student questionnaire.**

From an analysis of the response patterns to items in the questionnaire students indicated that the assessment was relatively easy to complete, that they typically had done the range of applications indicated, that they were fairly confident in using ICT, and had a high level of ICT skills across the range of types of applications. On average they indicated using ICT for about 1.75 hours per day at school with the lowest 16 minutes and the highest 468 minutes.
At least half of the students had either done the nominated activities in Q10 or were confident they could do them with 93% having typed an assignment for school. Most (<75%) liked using computers at school and home but half only sometimes liked finding out things for themselves. Over 80% indicated great confidence in using computers.

All had Broadband Internet access at home and 84% used computers on a daily basis at home.

Almost all Strongly Agreed or Agreed with Q2 items concerning value of eAssessment (only 2 disagreed on any of these items). Most had some experience with tests on computers over a half believing they could quickly get used to it, but three (7%) felt that it could take them a long time.

The open-ended response items asked students to list up to three things they liked about doing the examination and three things they did not like. Generally students considered that using computers made it easier and was fun, and also provided a better environment within which they could use their skills and demonstrate their ideas.

The main concerns were about time management (particularly because they could access so many tools and so much information), poor typing skills, the distraction of others around with a potential to cheat and that the computer could crash (this only occurred for one student). Eight students could find no faults.

**Students Forum**

The students’ interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.

The students were very positive about doing the computer-based exam. They quite liked the idea that it was hands-on and they preferred it to the other assessments they had done. They thought other students were also positive toward doing the exam. They thought they were able to do good work and the IT worked well for them.

The things that they wanted to change in the future were, that they thought they needed more instructions and more details, particularly of what was specifically required. They gave examples… in that it didn’t say that graphics manipulation was important, or what actual skills were looked for in the assessment, and they would have liked more or direction on the things they needed to include. This might be because in previous tasks they had more specific direction.

They could not identify any technical problems in doing the exam. They were however, concerned that the layout of the classroom meant that they had to sit very close to each other, which they felt was inappropriate for an exam situation because they could see each other’s work. Other than that, they were very positive toward the exam and couldn’t think of anything else that needed improving for the exam.

**Structured Interview with Teacher**

The teacher believed that the task suited both his needs as a genuine assessment tool and the administration’s requirement for a Semester Exam. He felt that the delivery had worked well and there were no significant problems, even for the one class that was rushed for time due to an assembly. He was confident that the students preferred an electronic based exam over a written one as they thought that this gave them a better chance of demonstrating what they knew and what they could do. In particular they were keen to see how they compared with other students in what they could do with the technology. However, he believed that due to pressures of time and the exam, the students had produced lower quality work than expected, although some with poor literacy had done better in the graphics environment.

He believed that this type of exam should be an essential component of the course although he was concerned that the external assessment may only be paper-based. He believed that the only major
improvement that could be considered was to make the exam more difficult to increase the pressure on the students to perform and differentiate between them more.

Results of Marking

The exam output for 46 students (3 students did not provide work) were collected on a USB Flash Drive, and copied from the school’s server upon which the students saved their work. Each student had a folder containing at least one file that was the poster. The formats of the files varied depending on the software that they had used. Before marking could commence all the posters were transformed into PDF files with this file and the original file being uploaded into the MAPS online portfolio system.

The students’ work was firstly marked by two external assessors using a standards-based rubric, and then by five external assessors using a comparative pairs method of marking. At the same time the teacher marked the students’ work using his own analytic marks-based system. The two external assessors marked the student work on the three criteria associated with Outcome 3 of the course while the teacher marked using analytical marks and converting these to levels for Outcome 1 and 2. The assessor judgements on the outcome was converted to numbers (ranging from 12 to 25) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The student work was also assessed by five assessors using the paired comparisons method and the same three criteria associated with the course outcome.

For the standards-based marking method the assessors used the rubric provided in the course document based on Outcome 3, as described below. For the overall outcome and each ‘Aspect’ of the outcome, levels of achievement are described for Levels 4 through to 8.

Outcome 3 - Quality of Information Solutions: Students explore alternatives and use skills, techniques, processes, standards and conventions to achieve information solutions.

Aspect 1: apply accepted standards and conventions to create or modify various forms of information solutions;

Aspect 2: apply efficient skills, techniques and processes in the use of ICT systems to develop information solutions; and

Aspect 3: apply enterprising capabilities while exploring alternatives and working to achieve information solutions.

The assessors were given the following instructions on the three criteria to assess before making an overall assessment.

Criteria 1: The implied message and purpose of the poster and techniques used to deliver to the appropriate audience. Standards and conventions that may apply are the use of space, positioning of objects relative to each other, attraction when printed and hung, L to R and T to B convention, and use of text and graphics attributes to convey message(s).

Criteria 2: Use of Powerpoint, Publisher, Word, Corel PhotoPaint, Fireworks or CorelDraw will be an indication of technical skill. Quality of image combination and editing. Use of overlays, transparency etc.

Criteria 3: Innovation displayed in design layout or techniques. Inventiveness in using the required graphic images and text to convey a message.

For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Average Assessors</th>
<th>Pairs (%)</th>
<th>Teacher (%)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>XA101</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>60.3</td>
<td>75</td>
<td>66.5</td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.20: Tabulation of results from marking.
A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

![Correlation Table]

Figure 4.21: Correlation coefficients from the marking of the student work.

The external markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.70 (p<0.0001) on the ranking of students. This increased with some discussion of the results. Initially they disagreed substantially on XA121, XA126, XA143, XA145, XA148 (L4/5) and XA122 (L4 or below) of the 46 students.

The comparative pairs method of marking was also used with five assessors each making 420 comparisons using a digital marking tool. This provided a reliable set of scores (separation index of 0.973) with rankings highly correlated to that of the two assessors using the standards framework.

The teacher used a numerical marks allocated to each component of the exam to provide a mark out of 100 for each student that was converted to a level. The teacher also provided the semester total percentage for each student.

Comparison with external marking results indicated only limited agreement. When compared with the two assessors the correlations on the ranking of students were 0.20 and 0.33 respectively, with only the latter significant at the 0.05 level of significance. There was a 0.39 correlation coefficient with the
comparative pairs results, significant at the 0.05 level. The students’ semester marks ranking was significantly correlated with all the other measures at least at the 0.05 level.

**Comparative Pairs Marking and Analysis**

There were three further types of analysis conducted using the comparative pairs marking results. The first was to determine cut scores for Level 5 and Level 4 based on the standards-based rubric for the assessed outcome. The second addressed the question concerning holistic marking compared with weighted criteria marking. The third was to investigate the minimum number of comparisons the five judges could have made and still gaining an acceptably reliable measure.

**Cut Scores**

The comparative pairs marking analysis provided a ranking for the students based on their location scores. The two assessors involved in the standards-based marking then determined the most likely cut scores on the holistic judgements for Level 5 and Level 4 by looking back at their marking notes, looking at a number of pieces of work again, and determining reasonable gaps in the location scores. As a result seven students were allocated Level 5, 29 students Level 4, and 10 students Level 3.

In comparing the results from the standards-based marking and the comparative pairs marking it was found that a few students would be ranked significantly differently for the two methods. In particular one student who the two assessors judged to be at a Level 5 for the Outcome using the standards-based rubric finished with scores from the comparative pairs marking that would clearly give a Level 4, similarly for four others judged by at least one assessor to be at a Level 5, and two judged to be below Level 4 would have received a Level 4 and four judged at Level 4 would have received a Level 3.

It was felt that to some extent the discrepancy in the results gained by these individuals students through the two methods of marking was due to differences in the weighting given to the criteria by assessors in the comparative pairs holistic marking. In particular, it is likely that the weighting given to technical skill was considerably different with less emphasis on technical skill in the holistic marking. It was noted that there was a high correlation between the results from the comparative pairs marking for the holistic judgements and judgements on the first criteria. Therefore it was determined to investigate the weighting of the three criteria.

**Weighted Criteria**

When making their comparisons between pairs of student work assessors were asked to make four judgements, to select the better on each of three criteria and then select the better in a holistic sense. The three criteria concerned: (1) the extent to which the poster employed standards and conventions to convey information; (2) the level of technical skill applied; and (3) the extent to which innovation, creativity and individuality were displayed. The holistic judgement took account of these three criteria but fundamentally concerned how well the poster would ‘work’. A correlation analysis between the three criteria scores and the holistic (global) score is shown in the table below. All four scores were highly correlated ($p<0.0001$) with the highest correlation between the first criteria (A1) and the holistic score with the second criteria (A2) the lowest correlated. This probably is consistent with assessors tending to judge the posters more on the use of standards and conventions for posters than the level of technical skills exhibited.
The researcher decided to compare the holistic judgement with the result of combining the scores from the three criteria weighted 30%, 50% and 20% respectively. Two methods of combining the scores was investigated, firstly applying a multiplier to each criteria to put each on a scale from 0 to 100 before weighting them, and secondly converting the criteria scores to standard scores before weighting them. The table below shows the results of a correlation analysis between these two methods of combining the criteria scores (Weighted is the result of the first method and TOT_STSC the result of the second method) and the holistic (GLOBAL) score. In all cases the correlations were very high (p<0.0001) with almost perfect correlation between the two methods.

### Correlations

<table>
<thead>
<tr>
<th></th>
<th>A1 Pearson Correlation</th>
<th>A2 Pearson Correlation</th>
<th>A3 Pearson Correlation</th>
<th>GLOBAL Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.623*</td>
<td>.736*</td>
<td>.936*</td>
<td>.953*</td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

### Correlations

<table>
<thead>
<tr>
<th></th>
<th>WEIGHTED Pearson Correlation</th>
<th>Total Pearson Correlation</th>
<th>GLOBAL Pearson Correlation</th>
<th>TOT_STSC Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>1.000*</td>
<td>.949*</td>
<td>.998*</td>
<td>.998*</td>
</tr>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

The results were now investigated for the individual students for whom discrepancies had been previously determined. The one student who the two assessors judged to be at a Level 5 using the standards-based rubric but who finished with a Level 4 score from the holistic comparative pairs marking had improved from a score of 50 to 59 but still would not have received a Level 5. Of the other four judged by at least one assessor to be at a Level 5 three had changed little but the other had improved from 38 to 56. There had been little change to the scores of the two judged to be below Level 4 who would have received a Level 4. Three of the four judged at Level 4 who would have received a Level 3 had substantially improved scores that would have now afforded them a Level 4. Overall the results of the weighting were somewhat inconclusive when considering the impact on the outcome for some of the individual students, however, the use of weighting of criteria scores is perhaps more educationally defensible than the use of only holistic judgements.
Number of Comparisons
An analysis was conducted to check if reliability of the paired comparison judgments would significantly be reduced if the number of comparisons were reduced with effectively each assessor making 210 comparisons, and if the script locations would shift significantly. Since in each pair of two exemplars in the pairwise comparison judgments were compared twice, it was possible to split the two pairs and still have each exemplar compared with each other. The authors then ran the Rasch paired comparison analysis for each split data set and examined whether the exemplar locations shifted significantly. Table 4.9 below shows the location of each exemplar from the global judgements for the two split groups.

Table 4.9  
Item location, Standard Errors for Group 1 and Group 2 for global judgements.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exemplar</td>
<td>Location</td>
</tr>
<tr>
<td>34</td>
<td>4.08</td>
</tr>
<tr>
<td>3</td>
<td>4.017</td>
</tr>
<tr>
<td>33</td>
<td>3.709</td>
</tr>
<tr>
<td>28</td>
<td>3.108</td>
</tr>
<tr>
<td>40</td>
<td>2.875</td>
</tr>
<tr>
<td>26</td>
<td>2.845</td>
</tr>
<tr>
<td>5</td>
<td>1.287</td>
</tr>
<tr>
<td>30</td>
<td>2.624</td>
</tr>
<tr>
<td>11</td>
<td>2.06</td>
</tr>
<tr>
<td>34</td>
<td>2.077</td>
</tr>
<tr>
<td>27</td>
<td>1.793</td>
</tr>
<tr>
<td>15</td>
<td>1.424</td>
</tr>
<tr>
<td>45</td>
<td>1.629</td>
</tr>
<tr>
<td>41</td>
<td>1.094</td>
</tr>
<tr>
<td>13</td>
<td>1.08</td>
</tr>
<tr>
<td>1</td>
<td>1.072</td>
</tr>
<tr>
<td>46</td>
<td>0.764</td>
</tr>
<tr>
<td>36</td>
<td>0.758</td>
</tr>
<tr>
<td>37</td>
<td>0.596</td>
</tr>
<tr>
<td>7</td>
<td>0.588</td>
</tr>
<tr>
<td>24</td>
<td>0.268</td>
</tr>
<tr>
<td>23</td>
<td>0.267</td>
</tr>
<tr>
<td>8</td>
<td>0.115</td>
</tr>
<tr>
<td>38</td>
<td>0.103</td>
</tr>
<tr>
<td>2</td>
<td>-0.044</td>
</tr>
<tr>
<td>43</td>
<td>-0.055</td>
</tr>
<tr>
<td>12</td>
<td>-0.204</td>
</tr>
<tr>
<td>35</td>
<td>-0.22</td>
</tr>
<tr>
<td>42</td>
<td>-0.548</td>
</tr>
<tr>
<td>25</td>
<td>-0.719</td>
</tr>
<tr>
<td>39</td>
<td>-0.72</td>
</tr>
<tr>
<td>9</td>
<td>-0.875</td>
</tr>
<tr>
<td>17</td>
<td>-1.614</td>
</tr>
<tr>
<td>29</td>
<td>-1.635</td>
</tr>
<tr>
<td>32</td>
<td>-2.045</td>
</tr>
<tr>
<td>6</td>
<td>-2.413</td>
</tr>
<tr>
<td>16</td>
<td>-2.433</td>
</tr>
<tr>
<td>18</td>
<td>-2.443</td>
</tr>
<tr>
<td>19</td>
<td>-2.676</td>
</tr>
<tr>
<td>20</td>
<td>-2.683</td>
</tr>
<tr>
<td>21</td>
<td>-2.929</td>
</tr>
<tr>
<td>22</td>
<td>-3.191</td>
</tr>
<tr>
<td>4</td>
<td>-3.691</td>
</tr>
<tr>
<td>10</td>
<td>-3.715</td>
</tr>
<tr>
<td>14</td>
<td>-3.729</td>
</tr>
<tr>
<td>31</td>
<td>-3.786</td>
</tr>
</tbody>
</table>
NB: In Table 4.9 the scripts are ranked from highest (most number of times preferred/chosen) to lowest (least number of times preferred). The Location is the value assigned to the script in logits (logarithmic units of measurement). The std err is the standard error of measurement (SEM) produced by the (Rasch) model calculated from the reliability (r) and the standard deviation of scores (S_x) of an examination using the relationship

\[ SEM = S_x \sqrt{1 - r} \]

The score is the actual number of times the script was preferred in all the comparisons in which that script was involved.

Some item locations significantly change with different judges. For example, in the global judgement, exemplar number 33 could have been rated on different criteria by the two raters because this was a global judgement. In their rating, judges may have emphasised certain aspects of their global judgement to be stronger. This may support the view that, in high stakes assessment, one rater should not be used to determine the location of two exemplars.

Figure 4.24 shows the correlation of exemplar locations between the two split comparisons between the two groups in the global comparison. The results show that the ratings provided by Group 1 raters and Group 2 raters was highly correlated.

<table>
<thead>
<tr>
<th>Location group 1</th>
<th>Location group 2</th>
</tr>
</thead>
</table>
| Pearson Correlation | 1 | .982(***)
| Sig. (2-tailed)  | .000 |
| N                | 46 | 46 |
| Location group 2 | Pearson Correlation | .982(***) | 1 |
| Sig. (2-tailed)  | .000 |
| N                | 46 | 46 |

* Correlation is significant at the 0.01 level (2-tailed).

Figure 4.24: Correlation of Group 1 and Group two for Global judgements

Table 4.10 below shows the reliability in terms of Separation Index from comparisons based on 2100 comparisons (when all exemplars are rated at least twice, by two different markers) and 1035 comparison (when all exemplars are rated only once). The Separation Index (SI) performs a similar function to Cronbach’s Alpha, a measure of internal consistency (reliability) but SI has the advantage that it can be used when there is missing data, which is often the case in educational research.

Table 4.10
Reliability in terms of separation index when number of comparisons is reduced.

<table>
<thead>
<tr>
<th>Type of Judgement</th>
<th>SI for 2100 comparisons of 46 exemplars</th>
<th>SI for 1035 comparisons of 46 exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect 1</td>
<td>0.973</td>
<td>0.963</td>
</tr>
<tr>
<td>Aspect 2</td>
<td>0.960</td>
<td>0.928</td>
</tr>
<tr>
<td>Aspect 3</td>
<td>0.955</td>
<td>0.931</td>
</tr>
<tr>
<td>Global</td>
<td>0.973</td>
<td>0.955</td>
</tr>
</tbody>
</table>

Table 4.10 above shows that when the numbers of comparison are reduced, as long as all exemplars are compared with each other at least once, the Separation Index remains high. Reduction in the number of comparisons reduces the Separation Index only marginally and the reduction is not considered statistically significant.
**Conclusions**

Overall it was concluded that this computer-supported production exam was implemented with no significant technical difficulties with all students able to complete the requirements of the exam in the time permitted.

The students responded well to the style of exam and generally preferred it to a theory paper-based exam. Capturing the performance of the students was relatively straightforward using the school server copied to a USB drive and then converting all output to PDF files for uploading with the original file into the MAPS portfolio system.

The structure of the assessment task made it difficult for students to demonstrate high level achievement and difficult for assessors to recognise that achievement. This was particularly due to the lack of student annotation or response to indicate their intentions in the design. It would have been useful to have 15 minutes for annotations at the end for students to explain what they did and why (information design and techniques).

In terms of the marking of the student work the following conclusions could be drawn.

1. There were clear differences in results based on marking by outcomes using the standards framework compared with analytic numeric scoring.
2. There was a high level of agreement between the two external assessors using the standards framework.
3. Only nine students were judged to have demonstrated any work above Level 4 and eight students were judged by at least one assessor to not have demonstrated work at even Level 4 for at least one outcome.
4. The comparative pairs method of marking provided a highly reliable set of scores that were strongly correlated to those of the assessors.

The investigations into the weighting and combining of scores on component criteria compared with scores from holistic judgements were successful. The comparative pairs scores on the component criteria were equally as reliable as the holistic scores but correlated to varying degrees with the holistic scores. When different weightings were applied to the three component criteria scores before summing to provide a total score that provided some differences in ranking from the holistic scores. Although inconclusive, for some students, this provided a better match to the standards-based rubric outcome level statements than for rankings from the holistic scores. While this was an interesting result it clearly needs more research.
5 - Engineering Case Studies

There were six schools, six teachers and six classes of Year 11 students involved in the project associated with the Engineering course. Six case studies were constructed and are reported upon in this section. The case studies are briefly introduced below.

<table>
<thead>
<tr>
<th>Case</th>
<th>Type of Assessment</th>
<th>Classes and Courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>Portfolio: vehicle design (one semester)</td>
<td>Year 11 – Eng 2A/B</td>
</tr>
<tr>
<td>ME</td>
<td>Production exam (one semester)</td>
<td>Year 11 – Eng 2A/B</td>
</tr>
<tr>
<td>VE</td>
<td>Performance tasks (four hours)</td>
<td>Year 11 – Eng 2A/B</td>
</tr>
<tr>
<td>WE</td>
<td>Production exam (one semester)</td>
<td>Year 12 – Eng 3A/B</td>
</tr>
<tr>
<td>HE</td>
<td>Portfolio (one semester)</td>
<td>Year 11 – Eng 2A/B</td>
</tr>
<tr>
<td>YE</td>
<td>Performance tasks (one semester)</td>
<td>Year 11 – Eng 2A/B</td>
</tr>
</tbody>
</table>
GE: Research portfolio

The GE case study involved one class of Year 11 students at one school completing the Engineering Studies course Units 2A-2B. There were 14 students in the class that was in a vehicle design context.

The aim was for each student to collate a portfolio of work, stored in digital formats, that was a part of a larger go-cart design and production task. It was to reflect the student’s initial research into aspects of go-cart design including safety, ergonomics and anthropometrics, market research, social, environmental and fuel economy considerations, steering mechanisms and safety. The assessment addressed Outcomes 1 and 4 of the Engineering course and used the related levels of achievement.

The teacher assessed the work for the purposes of student grades. Two external assessors assessed the work as part of the research being undertaken. This assessment did not influence the assessment of the teacher but was compared and analysed. The teacher developed the task based on the previous years work, and the accompanying portfolio requirements based on materials provided by the CC.

Students were to complete their portfolios in Term 2, preparatory to beginning work on their go-cart design and manufacture the following term. Some students completed their portfolio research early and began their specific design work and others were still completing their portfolio when workshop practice had begun. It was the intent of the teacher for the students to begin some fabrication in the workshop, and then record their practice in order to provide evidence of adhering to health and safety requirements. Not all students progressed to this stage. The portfolios were all developed in PowerPoint and included a range of sources of media including scanned sketches and digital pictures. In order to achieve the outcomes students were instructed to develop a portfolio which included aspects of market research, safety, ergonomics and anthropometrics, steering mechanisms, engineering and society and the environment in relation to vehicles generally and go-carts specifically.

Results from data analysis

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Survey of Students

There were 13 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.

Table 5.1

<table>
<thead>
<tr>
<th>Description</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
<td>13</td>
<td>2.8</td>
<td>3.8</td>
<td>3.3</td>
<td>0.312</td>
</tr>
<tr>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
<td>13</td>
<td>1.6</td>
<td>3.0</td>
<td>2.3</td>
<td>0.388</td>
</tr>
<tr>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
<td>13</td>
<td>1.8</td>
<td>2.8</td>
<td>2.3</td>
<td>0.25</td>
</tr>
<tr>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
<td>13</td>
<td>2.2</td>
<td>2.7</td>
<td>2.4</td>
<td>0.20</td>
</tr>
<tr>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
<td>11</td>
<td>2.4</td>
<td>3.7</td>
<td>3.1</td>
<td>0.56</td>
</tr>
<tr>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
<td>13</td>
<td>16</td>
<td>192</td>
<td>82.4</td>
<td>61.84</td>
</tr>
</tbody>
</table>
Few students had done an examination on a computer before and felt they would need some time to get used to the process. They felt doing the test on the computer was quick and easy, and it was an appropriate tool to use in the development of the portfolio. All the students except one had all the technologies listed in Q5 at home, and all had broadband Internet at home. Most students used their computers most days, and all more than once a week. At school they used computer for an average of 1 hr 20 minutes per week. Most students were confident with computers and felt they were good at using them. Of the types of computer software and equipment listed in Q13, students felt least confident about databases and spreadsheets.

**Students Forum**

The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the exam; and any other thoughts they had about doing exams on computers.

The students explained that they were basically doing a design process for a go-cart, planning out everything, gathering up the research and when working in the workshop taking photos and video. They put it into a PowerPoint and some made a design brief. They put all information into a PowerPoint. The whole assignment was done in PowerPoint and they put all the parts together for it. They found the process “quite good” as it was “easy to change your work”. One student commented that he found when he put the information into PowerPoint rather than Word the layout was better. They found the project was quite structured. They indicated that they did pre-design research like environmental factors, the construction, making the steering and all the parts by a certain date.

They believed that such assessment was “good in engineering” because they could show what they were doing, what processes they were using and they often added visual aids, like in describing the brake system. They added information on safety equipment and the like and found PowerPoint good for that. It was easy to show and share their ideas in front of the classroom unlike for a Word document. They also believed that because they were using PowerPoint it looked better. They could scan and print out, as all the programs they needed were accessible all the time on the network. The applications were accessible anywhere in the school such as the library. They found it useful using a computer to store drawings and other information. They also found it fast and easy to mix music and put into PowerPoint. They felt that it was much faster to write an electronic portfolio.

Overall they liked the assessment and did not want any major changes. One student thought that there should be more manufacturing but conceded that “some things cannot be manufactured”. Some felt that there was difficulty getting some parts in. They pointed out that they all got the same kit (e.g. roll cage and common engines) but they designed things like what kind of brakes they wanted to use.
The students couldn’t identify many technical problems. They indicated that some students had problems with the scanner and inserting into PowerPoint. They scanned the images and emailed them to themselves because they didn’t want to crash the computer or lock the system but simply scan their work. As the video file sizes were very big when they used the multimedia video software they would send it in parts instead of the whole PowerPoint. As the file sizes were very big they couldn’t send via the school’s email server so one student emailed from home so it was easy to hand in. They explained that when the teacher marks their work he usually writes comments about their ideas or correction on the work and sends it back. They had no other problems as they were a group of 14 students. They felt that it was “quite a good program”. The only problem was they said that it hadn’t been examined yet and they hadn’t been told when it was happening. They would like to see this kind of task expanded.

Results of Marking

The portfolio output for 14 students were collected on a DVD, copied from the school’s server upon which the students saved their work. Each student had a folder containing folders and files that made up their portfolio. The formats of the files included HTML, DOC, PUB, and PPT. Before marking commenced all portfolios were checked, zipped and uploaded into the MAPS online portfolio system.

The students’ work was marked by two external assessors using a standards-based rubric. At the same time the teacher marked the students’ work using his own analytic marks-based system. The two external assessors marked the student work on the criteria associated with the four outcomes of the course while the teacher awarded a level for each outcome that was converted to mark out of 100 with 25 allocated for a level 4 for each outcome. The assessor judgements on the outcome were converted to numbers (ranging from 12 to 25) using the Curriculum Council’s converter with the lowest Level 4 equating with a score of 12.

For the standards-based marking method the assessors used the rubric provided in the course document based on Outcomes 1 and 4, as described below. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the following instructions on the criteria to assess.

<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Engineering Process: Students apply and communicate a process to design, make, and evaluate components and systems.</td>
<td></td>
</tr>
<tr>
<td>1.1 Identify needs and opportunities (L4)</td>
<td>Consider the appropriateness of a range of technologies (L5) Justifies the appropriateness of the selected technology (L6)</td>
</tr>
<tr>
<td>1.2 provides a range of optional proposals supported with graphics (L4)</td>
<td>Justifies the option selected (L5) Incorporates appropriate technological language (L6)</td>
</tr>
<tr>
<td>4 - Engineering in Society: Students understand the interrelationships between engineering projects and society.</td>
<td></td>
</tr>
<tr>
<td>4.1 Identify community influences on engineering (L4)</td>
<td>Considers social factors (L5) Considers environmental implications of their proposal (L6)</td>
</tr>
<tr>
<td>4.2 identified the influence of engineering on society (L5)</td>
<td>Details the influence of a specific event or technology on society (L6)</td>
</tr>
</tbody>
</table>

The two external assessors marked the student work on the aspects of the two outcomes (2 & 4) while the teacher also marked using levels but in a more holistic manner. All the judgements on the two outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12. The numeric scores for each assessor were summed across the outcomes and averaged. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.
A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

**Correlations of Marks**

<table>
<thead>
<tr>
<th>(N=14)</th>
<th>TOT_M1</th>
<th>TOT_M2</th>
<th>AVG</th>
<th>TCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT_M1</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.540(*)</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.046</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOT_M2</td>
<td>Pearson Correlation</td>
<td>.540(*)</td>
<td>1</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.046</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tch</td>
<td>Pearson Correlation</td>
<td>.407</td>
<td>.555(*)</td>
<td>0.481</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.149</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).

**Correlations of Rankings**

<table>
<thead>
<tr>
<th>Rank of TOT_M1</th>
<th>Rank of TOT_M2</th>
<th>Rank AVG</th>
<th>Rank of Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td>.452</td>
<td>0.726</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.104</td>
<td></td>
<td>.203</td>
</tr>
<tr>
<td>Rank of TOT_M2</td>
<td>Pearson Correlation</td>
<td>.452</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.104</td>
<td></td>
<td>.064</td>
</tr>
<tr>
<td>Rank of Tch</td>
<td>Pearson Correlation</td>
<td>.362</td>
<td>.507</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.203</td>
<td>.064</td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).

The two external markers used the aspects and level statements for the two outcomes to mark the portfolios. The markers had a significant level of agreement with an initial inter-rater reliability correlation of 0.54 (p<0.05) on the marks of the students, but a correlation that was not significant on the rankings of the students.

The teacher marked the students work using outcomes and levels. These were then converted to a mark between 12 and 25 to represent levels between 4 and 8. Comparison with external marker results indicated little agreement apart from the top ranked students. When compared with the two assessors the inter-rater reliability correlations on the ranking of students were 0.362 and 0.507 respectively, with neither result statistically significant. The mean scores were quite different for all three markers possibly indicating a different understanding and interpretation of the standards at each level.

**Conclusions**

This research portfolio was a preparatory task to manufacturing a go-cart, consequently the students were keen to complete the research. They could see the relevance and need to do the preparatory work, and so were gratified that their examination was part of the work they had to do anyway, and not something extra that they had to prepare for, although some were waiting for a separate examination.

The students developed their portfolio in Powerpoint, and were happy with its flexibility and ease of use as it seems some of them had not used it before for this purpose. There were few technical problems with some scanning issues, and the task provided scope for a range of outcomes.

Marking was done on the aspect statements for each level, but were obviously interpreted in different ways by the assessors and teacher, given the lack of correlation between the results.
ME: Portfolio Based on Template

The ME case study involved one class of Year 11 students at one school completing the Engineering Studies course Unit 2B. There were 4 students in the class that was in a design context.

The aim was for each student to collate a portfolio of work, stored in digital formats representing their research and planning work for an Engineering project. The portfolio content centred around a design brief where the students would redesign an electronic toy explorer for the target group of ten to fifteen year olds whilst considering safety, materials, time, cost and gender. A template portfolio was provided to the students, which reflected the requirements of the design brief. The assessment addressed Outcomes 1, 2, 3 and 4, which were elaborated in terms of levels 4-8 of the Engineering course. The teacher assessed the work for the purposes of student grades. Two external assessors assessed the work as part of the research being undertaken. This assessment did not influence the assessment of the teacher but was compared and analysed. The teacher developed the tasks and portfolio template based on materials provided by the Curriculum Council.

Students were required to progressively work on their portfolio throughout the semester as they worked through the requirements of the design task. These requirements addressed aspects of each of the outcomes that were appropriate to the task. The teacher provided workshop time specifically for working on their portfolio.

Results from data analysis

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class

Students were working on their portfolios throughout the semester, with some specific class time set aside for portfolio work, and other planning and research work being done when needed by the student in order to progress. The prevailing workshop culture was for students to minimise their research and documentation work and to maximise their practical workshop activities.

Survey of Students

While there were only 4 students in the class that had their portfolio assessed there were 9 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.

Table 5.2

<table>
<thead>
<tr>
<th>Description</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
<td>9</td>
<td>2.4</td>
<td>3.6</td>
<td>3.1</td>
<td>0.39</td>
</tr>
<tr>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
<td>9</td>
<td>1.8</td>
<td>3.0</td>
<td>2.3</td>
<td>0.51</td>
</tr>
<tr>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
<td>9</td>
<td>1.8</td>
<td>2.6</td>
<td>2.3</td>
<td>0.30</td>
</tr>
<tr>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
<td>9</td>
<td>2.3</td>
<td>2.7</td>
<td>2.5</td>
<td>0.12</td>
</tr>
<tr>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
<td>9</td>
<td>2.7</td>
<td>4.0</td>
<td>3.5</td>
<td>0.47</td>
</tr>
<tr>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
<td>9</td>
<td>0</td>
<td>66</td>
<td>20.11</td>
<td>20.34</td>
</tr>
</tbody>
</table>
From the table above it appears that while the students had not generally done a portfolio examination in a workshop before, they had little difficulty with the task and the integration of computers into the workshop design task. While less felt that the computer was a useful tool for reflection, and that the electronic portfolio adequately reflected their capability, they felt that using computers made school-work easier. They were familiar with computers and related technology, all having a computer and Internet access at home, and using it almost every day. They had a positive attitude toward computers and feel confident in using them and that they are easy to use. They feel quite skilled with a broad range of computer activities.

**Students Forum**

The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the task; and any other thoughts they had about doing assessment on computers.

The students commented that they had found the task quite interesting and one reason for this was that it was different to what they had done in engineering in previous years. Other reasons the students gave for the extra interest in the task were that usually they only do the design for the craft but this time they did the programming as well and that they had not utilised the microchips before so the experience was novel.

They found the emphasis on making things preferable to only completing a portfolio, principally because creating a portfolio mainly consists of theoretical design drawings as opposed to the practical aspect, which they found easier. If making things took too long however, they were put aside so the portfolio portion could be completed. One student commented that it was advantageous that he had done programming last year because the programming was quite challenging.

The students appreciated the variety that this task bought and commented “last year it was easy – you wrote and drew the body in a CAD program, you built it and raced it. That was the whole task. Everyone made the same thing”. The students also enjoyed using “more advanced programs this time” and they made the point that they needed longer sessions, as they could not get enough work done in a 45-minute period.

The student encountered some difficulties with soldering and incompatibilities with displaying graphics when using Macintoshes and Windows machines. They reported a further problem with having to scroll up and down looking for information when they were doing their test. They appreciated having the Internet available during the test as they could use a dictionary when needed.
Teacher Structured Interview

A structured ‘interview’ was carried out via email with the teacher employing the task in the school. The questions forming the interview covered what they thought of the task and its structure and potential, how the students reacted to the activity and the quality of their output, and whether there were any technical or other problems with the implementation.

Overall, the teacher was very please with the whole outcome of the task. They felt that the task of producing a digital portfolio using a template that the teacher had created scaffolded the weaker students while giving those with more confidence a base to build upon. They believed the structure of the activity meant that the students took ownership of the task from the outset. The sub-tasks and instructions went well and the students produced good quality work. However, some of the less-able students had difficulty keeping up due to the shortness of time. There were not any technical issues related directly to the task although there were some problems with the computers being used. The teacher felt that some of the students needed some prior training with the software required for the activity.

Results of Marking

Two external assessors and the teacher marked the 13 students’ work using a standards-based rubric. The teacher designed the level statements for the rubric. All the judgements on the outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12. The numeric scores for each assessor were averaged across the outcomes.

For the standards-based marking method the assessors used the rubric provided in the course document based on Outcomes 1 to 4, as described below. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given a rubric of instructions on the criteria to assess (see below).

For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Average (Avg)</th>
<th>Teacher (Tch)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME101</td>
<td>16.3</td>
<td>16.3</td>
<td>16.3</td>
<td>20.5</td>
<td>2.0</td>
</tr>
<tr>
<td>ME102</td>
<td>17.3</td>
<td>16.8</td>
<td>17.05</td>
<td>18.5</td>
<td>1.0</td>
</tr>
<tr>
<td>etc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

*Figure 5.5: Tabulation of results from marking.*
A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

**Score Correlations (N=4)**

<table>
<thead>
<tr>
<th></th>
<th>TOT_M1</th>
<th>TOT_M2</th>
<th>AVG</th>
<th>Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT_M1</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.992(**)</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.008</td>
<td></td>
<td>0.342</td>
</tr>
<tr>
<td>TOT_M2</td>
<td>Pearson Correlation</td>
<td>.992(**)</td>
<td>1</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.008</td>
<td></td>
<td>0.256</td>
</tr>
<tr>
<td>Tch</td>
<td>Pearson Correlation</td>
<td>.658</td>
<td>.744</td>
<td>0.701</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.342</td>
<td>.256</td>
<td></td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

**Rank Correlations**

<table>
<thead>
<tr>
<th>Rank of</th>
<th>Rank of</th>
<th>Rank AVG</th>
<th>Rank of Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT_M1</td>
<td>TOT_M1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td>1.000(**)</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
<td>.200</td>
</tr>
<tr>
<td>TOT_M2</td>
<td>TOT_M2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1.000(**)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
<td>.200</td>
</tr>
<tr>
<td>Tch</td>
<td>Tch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.800</td>
<td>.800</td>
<td>0.800</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.200</td>
<td>.200</td>
<td></td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

The two external markers used the level statements for the outcomes that had been designed by the teacher to mark the exam outputs. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.99 (p<0.01) on the marks given to the students and a correlation of 1 on the ranking of the students.

The teacher used the level statements that he had designed for the three outcomes to provide a mark out of 100 for each student, which was then converted to a mark out of 25. Comparison with external markers results indicated a non-significant correlation of 0.658, and a rank correlation of 0.8, which is high but again not significant (due to the low N). Agreement was evident in the rankings of the lower two students and the upper two students.

**Conclusions**

Overall it was concluded that this electronic portfolio exam was implemented with no significant technical difficulties with all students able to complete the requirements of the portfolio exam in the semester. Some students could have achieved a better grade had they spent more time on aspects of their portfolio, but that is often inevitably the case. The students responded well to the portfolio style of exam and generally preferred it to a theory paper-based exam although some felt it did not fully reflect their capacity and ability. Capturing the performance of the students was relatively straightforward, although some students expressed frustration in working on a different platform (PC) than the teacher (Mac) with the consequent confusion of collected data.

The portfolio template may have had a constraining effect on student performance, with spaces provided being indicative of the extent of the research, and the provided questions for analysis forcing students to address those questions regardless of the nature of their personal design direction.

In terms of the marking of the student work the following conclusions could be drawn.

- There was a high level of agreement between the two external assessors using the standards framework.
- The teacher was more generous with the marks given to students resulting in a higher group mean. The teacher may have been aware of aspects of individual student capability that was not fully reflected in the portfolio.
**VE: Focussed performance task**

The VE case study involved one class of Year 11 students at one school completing the Engineering Studies course (Unit 2A). There were 4 groups of 3 or 4 students in one class working in a mechanical fabrication context.

The aim was for students to work in groups of 3 or 4 to solve a practical design problem and document their progress and decision-making. The scenario was to design and produce a childproof automatic locking mechanism for a swimming pool gate. The form of documentation was up to the students, but all groups concluded with a PowerPoint based portfolio. The activity was conducted over two two-hour sessions on different days. Assessment judgement was based on the portfolio and the final product and addressed all four of the Engineering outcomes and used the related levels of achievement. The teacher assessed the work for the purposes of student grades while two external assessors assessed the work as part of the research being undertaken. This external assessment did not influence the assessment of the teacher but the two were compared and analysed. The teacher developed the task based on the previous years work, and the accompanying portfolio requirements based on materials provided by the Curriculum Council.

The students were divided into groups and presented with the task scenario. They were presented with a one-page outline of the scenario, the evidence expected in the portfolio, the suggested format for the evidence, a timesheet and the assessment criteria. Suggestions were made to the groups about the division of tasks based on expertise, the need to concurrently document the decisions and progress made and the materials that were available for the product development.

Students had access to all the materials and equipment available in the workshop, and computers in an adjacent room. Special timetable adjustments were made to permit this class of students two 2 hour workshop blocks of time. Each group had a camera to record their activities and a USB drive for submission of their portfolios. After the initial design stage of the project, each group was videotaped presenting their design ideas and reacting to questions and criticism from the other groups. This video became part of the electronic portfolio.

**Results from data analysis**

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

**Observations of the class**

The task was presented to the students toward the end of the semester as one of the forms of assessment for their semester grade. All students were quite intensively involved and felt the pressure to complete the task within the time allotted. At the beginning they were all involved in the initial design ideas development and then each worked on different sections of the product they were constructing, and one focused on developing the portfolio and documenting the group’s work.

**Survey of Students**

There were 9 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.
Table 5.3
Descriptions and descriptive statistics for the scales based on items from the student questionnaire (N=9).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>9</td>
<td>2.9</td>
<td>4.0</td>
<td>3.4</td>
<td>0.41</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>9</td>
<td>1.6</td>
<td>3.0</td>
<td>2.2</td>
<td>0.50</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>9</td>
<td>1.6</td>
<td>2.6</td>
<td>2.2</td>
<td>0.34</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confid</td>
<td>9</td>
<td>1.7</td>
<td>3.0</td>
<td>2.4</td>
<td>0.48</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>8</td>
<td>1.6</td>
<td>4.0</td>
<td>2.6</td>
<td>0.81</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>9</td>
<td>0</td>
<td>60</td>
<td>25.2</td>
<td>16.72</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

Some students had done practical examinations in a workshop before but felt they needed some more time to get used to it. They were all very positive about the ease of use and usefulness of the camera and computer in developing their portfolio.

All the students had all the digital equipment at home that was mentioned in Q5 except for the video camera and laptop, and all had either dial up or broadband Internet access. They use the computer at home more than once a week and spend about 25 minutes a day on a computer at school, and are quite confident about using computers for a range of activities.

Students Forum
The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the task; and any other thoughts they had about doing assessment on computers.

The students explained they were carrying out the design and construction of pool gate locking mechanism. They felt that doing an exam this was is better than doing a written exam and they enjoyed doing things around the class rather than sitting down with their books. None of the students had any trouble with the task also although most of them ran out of time. The students reported
creating a Powerpoint document of their portfolio for the project and they submitted this for marking. They said that they already understood how to use Powerpoint before the task began.

They believed that they did not need to worry about the exam beforehand and did not have anything to study. They did not suggest any future changes to the task, however, they had not yet received any feedback on their portfolio. During the actual test there were no reported technical problems although students occasionally had to wait for tools.

The students did not have any further thoughts on doing exams on computers.

**Teacher Structured Interview**

A structured ‘interview’ was carried out via email with the teacher employing the task in the school. The questions forming the interview covered what they thought of the task and its structure and potential, how the students reacted to the activity and the quality of their output, and whether there were any technical or other problems with the implementation.

The teacher believed that students thought that the task was well structured and that it kept them busy, interested and happy to be assessed. They commented that the students reacted very positively to this form of assessment preferring it to any other means of testing (particularly written exams) they had experienced in this course. In terms of results the students produced work of excellent quality for the time they had to complete the task.

Overall, the teacher felt that the assessment task worked well for Engineering Studies and had potential to be successful for assessment in other subject areas. They felt the task ran smoothly without any technical or other problems and the thorough prior set-up of the room assisted this.

**Results of Marking**

Two external assessors using a standards-based rubric marked the 9 students’ work. At the same time the teacher marked the students’ work using his own analytic marks-based system. The two external assessors marked the student work on criteria associated with the four outcomes of the course while the teacher also marked using levels but in a more holistic manner. The external assessors used the rubric based on the framework provided in the course document for the four outcomes, as described below. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the following instructions on the criteria to assess.
### Outcome/Aspect

#### 1 - Engineering Process: Students apply and communicate a process to design, make, and evaluate components and systems.

In achieving this outcome, students:
- investigate design needs and opportunities in engineering;
- generate engineering production proposals to solutions;
- manage engineering production processes to produce solutions and;
- evaluate intentions, plans and actions.

The student has clearly developed and defined design ideas for the problem;
The student had managed the engineering process.

#### 2. Engineering Understandings: Students understand properties of materials, energy transfer and design principles in engineering technologies.

In achieving this outcome, students:
- understand properties of materials and/or components in engineering technologies;
- understand energy transfer in engineering technologies; and
- understand design principles in engineering technologies.

The student provided a rationale for the selection of appropriate materials.

#### 3. Engineering technology skills: Students use materials, skills and technologies appropriate to the engineering industry.

In achieving this outcome, students:
- apply initiative and organisational skills;
- apply materials, techniques and technologies to achieve solutions to engineering challenges;
- operate equipment and resources safely; and
- apply skills of calculation and computation.

The student demonstrated group organisational skills; used equipment safely.

#### 4 - Engineering in Society: Students understand the interrelationships between engineering projects and society.

In achieving this outcome, students:
- understand how engineering technologies are influenced by beliefs and values; and
- understand beliefs and values are influenced by engineering technologies.

The student shows a clear understanding of social needs for this product.

The two external assessors marked the student work on the aspects of the four outcomes while the teacher also marked using levels but in a more holistic manner. The teacher awarded a level for each outcome that was converted to a mark out of 100 with 25 allocated for a level 4 for each outcome. For the external assessors all the judgements on the four outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12 and anything below that scoring 0. The numeric scores for each assessor were summed across the outcomes and averaged. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated and are shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Average (Avg)</th>
<th>Teacher (Tch)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE101</td>
<td>13.5</td>
<td>12.5</td>
<td>13.0</td>
<td>13.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below. The low correlations between the teacher and external markers maybe the result of
inadequately defined assessment criteria. The interpretation of outcomes aspects in terms of this project were quite brief which in turn allows more latitude in the awarding of marks.

### Correlations of Marks

<table>
<thead>
<tr>
<th></th>
<th>TOT_M1</th>
<th>TOT_M2</th>
<th>AVG</th>
<th>Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT_M1</td>
<td>Pearson Correlation</td>
<td>-3.59</td>
<td>3205</td>
<td>-775</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOT_M2</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>3205</td>
<td>.617</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.641</td>
<td></td>
<td>.383</td>
</tr>
<tr>
<td>Tch</td>
<td>Pearson Correlation</td>
<td>-7.75</td>
<td>.617</td>
<td>-.079</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.225</td>
<td>.383</td>
<td></td>
</tr>
</tbody>
</table>

### Correlations of Ranks

<table>
<thead>
<tr>
<th></th>
<th>Rank of TOT_M1</th>
<th>Rank of TOT_M2</th>
<th>AVG Rank</th>
<th>Rank of Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank of TOT_M1</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>-400</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank of TOT_M2</td>
<td>Pearson Correlation</td>
<td>-400</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.600</td>
<td></td>
<td>.225</td>
</tr>
<tr>
<td>Rank of Tch</td>
<td>Pearson Correlation</td>
<td>-7.75</td>
<td>.775</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.225</td>
<td>.225</td>
<td></td>
</tr>
</tbody>
</table>

### Conclusions

Overall it was concluded that this workshop exam and portfolio was implemented with no significant technical difficulties. The time allocation was a little short to permit comfortable completion of both the product and portfolio. The students enjoyed this type of examination, preferring it to a paper-based test. They felt it more adequately reflected their abilities, in addition to not having to study for it. Each group was able to divide up the tasks so that they were all busy for most of the time. The video tape of their group presentation was in a format that did not match the computers the students had access to, so it had to be handed in as a separate file rather than integrated into their portfolio.

All the students had the skills needed to digitally capture their performance and record it in a portfolio. This included scanning documents, taking still and video pictures and using Word and PowerPoint. The scenario to which they had to respond (locking mechanism for a swimming pool gate) was a context with which they were familiar and so they could respond in a workshop with minimal preparatory research.

In terms of the marking of the student work the following conclusions could be drawn.

- There were clear differences in the marks allocated by the teacher and the external assessors.
- Only one group was judged to have demonstrated any work above Level 4.
WE: Portfolio based on research

The WE case study involved one class of Year 12 students at one school completing the Engineering Studies course Unit 3A. There were 10 students in one class that was in a mechanisms and mechanical transfer context.

The aim was for each student to research a specific area of engineering (motion, mechanisms and mechanical transfer) and demonstrate their understandings in a portfolio. The portfolio was a component of a unit that comprised 6 tasks including a prototype and a product. Four of the six tasks in the unit comprised the final portfolio, the other two being the devising and development of a prototype. The students developed a range of ideas on lifting mechanisms, testing mechanical advantages such as using screw jack and a hydraulic press. They also examined changing motion. Based on that knowledge they had to apply these basic concepts in the next task to their own design.

There were three requirements in this task:

- Provide a description of a range of terms,
- Provide written and graphic evidence of force conversion,
- Use calculations to provide evidence of the mathematical relationships of energy transfer.

This assessment task addressed only Outcome 2 (aspects 1 and 2) and used the related levels of achievement. When the other 5 tasks in the unit were included, students were assessed on all four outcomes. The teacher assessed the work for the purposes of student grades. Two external assessors assessed the work as part of the research being undertaken. This assessment did not influence the assessment of the teacher but was able to be compared and analysed. The teacher developed the task and task requirements based on materials provided by the Curriculum Council.

Students were guided through the logically ordered tasks in order to develop the knowledge and competencies to produce and test a prototype. They were provided with some time in class to develop their understandings and produce their portfolio, but were also required to do some work outside of class.

Results from data analysis

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class

Students had access to the workshop and adjacent computer lab and, apart from class times when teacher instruction was taking place, the students could work in the area that matched their stage of development. The focus was on the development of understanding and compilation of evidence in a portfolio early in the semester, and then later in the semester the focus was more on the development and then testing of the prototype.

Survey of Students

There were 10 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.
Table 5.4

Descriptions and descriptive statistics for the scales based on items from the student questionnaire (N=10).

<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>10</td>
<td>2.8</td>
<td>3.8</td>
<td>3.2</td>
<td>0.27</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>10</td>
<td>1.4</td>
<td>2.8</td>
<td>2.0</td>
<td>0.39</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>10</td>
<td>1.4</td>
<td>2.6</td>
<td>2.1</td>
<td>0.37</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confidence</td>
<td>10</td>
<td>1.3</td>
<td>2.7</td>
<td>2.4</td>
<td>0.44</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>10</td>
<td>1.9</td>
<td>3.9</td>
<td>3.3</td>
<td>0.62</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>10</td>
<td>0</td>
<td>32</td>
<td>92.4</td>
<td>88.69</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

Figure 5.7: Distribution of scores for the Skills and SCuse scales on the student questionnaire.

From the survey it is clear that the students had little experience doing a test on a computer before and felt as though they needed some more time to get used to it. They felt comfortable using the IT and it suited the task, though some felt it had limited scope to demonstrate their full potential. Most students had the Internet and a range of IT equipment at home. Most used their home computer more than once a week, and spent an average of about 1½ hours each day on a computer at school. They were all positive about using computers and the role computers play. They were confident about using computers and positive about their skills in the range of areas described in Q13, though less positive about video editing and databases.

Students Forum

The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the task; and any other thoughts they had about doing assessment on computers.

The students felt that the concepts were good for them to learn but they didn’t like the report format or the written aspect. They also believed that there was too much Mathematics, especially ratios. The majority of the project was written whereas they would have preferred more of a 50/50 balance. They didn’t get into the workshop until week 8 meaning they only had 3 weeks in the workshop.
The students didn’t like the assignments particularly doing the written aspects. They felt it was nothing like they had done last year, which was more like a 50/50 balance between theory and practice. They complained that they already did so much writing in other subjects and that this project, with its large amount of written content to do in the design part, felt more like an English assignment.

The students found the limited Internet access frustrating when they cannot find information that gives solutions to their problems. They thought the single class was too short, getting an hour, which was basically 40 min (with 10 mins to get ready and 10 minutes to pack up). Some of machines were broken, and the mill was very inaccurate which was frustrating. Last year there were 4 people in the class but now there are 15 and students had to wait for other students to complete tasks on machines which cut down their effective work time. Therefore the students felt the classes should be smaller.

Students reported some technical problems with the Internet dropping out, however the school IT support was good so the problem didn’t last long. Other problems revolved around the students’ difficulty in understanding the questions and not being sure how long each task was to be.

Results of Marking

For the standards-based marking method the external assessors used the rubric provided in the course document based on Outcomes 2, as described below to mark the 10 students work. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the following instructions on the criteria to assess.

<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 - Engineering understandings: Students understand properties of materials and/or components, energy transfer and design principles in engineering technologies.</td>
<td>For example: understand relationships between properties and functions of materials, and/or components in circuits or systems, in engineering technologies (level 5)</td>
</tr>
<tr>
<td>Aspect: Understand properties of materials and/or components in engineering technologies</td>
<td>For example: understand different ways energy is converted (level 5)</td>
</tr>
<tr>
<td>Aspect: Understand energy transfer in engineering technologies</td>
<td>For example: understand different applications of design principles</td>
</tr>
<tr>
<td>Aspect: Understand design principles in engineering technologies</td>
<td></td>
</tr>
</tbody>
</table>

The two external assessors and the teacher marked the student work on the aspects of outcome 2 using a standards-based rubric. All the assessment judgements were initially on a level, and then were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12. For each assessor and for the teacher the students were ranked so that comparisons could be made. The teacher used the same level statements matrix for the outcome to give each student a level. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Average (Avg)</th>
<th>Teacher (Tch)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A1</td>
<td>A2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A1</td>
<td>A2</td>
</tr>
<tr>
<td>WE101</td>
<td>20.0</td>
<td>13.0</td>
<td>16.5</td>
<td>18.0</td>
<td>2.5</td>
</tr>
<tr>
<td>WE102</td>
<td>12.0</td>
<td>15.0</td>
<td>13.5</td>
<td>15.0</td>
<td>6.0</td>
</tr>
<tr>
<td>etc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.8: Tabulation of results from marking.
A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

### Correlations of Marks

<table>
<thead>
<tr>
<th>(N=10)</th>
<th>TOT_M1</th>
<th>TOT_M2</th>
<th>AVG</th>
<th>Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT_M1</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.006</td>
<td>0.503</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.991</td>
<td></td>
</tr>
<tr>
<td>TOT_M2</td>
<td>Pearson Correlation</td>
<td>.006</td>
<td>1</td>
<td>0.503</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.991</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tch</td>
<td>Pearson Correlation</td>
<td>.659</td>
<td>-.703</td>
<td>-.022</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.155</td>
<td>.119</td>
<td></td>
</tr>
</tbody>
</table>

### Correlations of Rankings

<table>
<thead>
<tr>
<th>(N=10)</th>
<th>Rank of TOT_M1</th>
<th>Rank of TOT_M2</th>
<th>Rank AVG</th>
<th>Rank of Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank of TOT_M1</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.029</td>
<td>.5145</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.956</td>
<td></td>
</tr>
<tr>
<td>Rank of TOT_M2</td>
<td>Pearson Correlation</td>
<td>.029</td>
<td>1</td>
<td>.5145</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.956</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank of Tch</td>
<td>Pearson Correlation</td>
<td>.537</td>
<td>-.702</td>
<td>-.0825</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.272</td>
<td>.120</td>
<td></td>
</tr>
</tbody>
</table>

Comparison with external marker results indicated little agreement, and even a complete reversal of rank positions with student WE106. Consequently the inter-rater reliability correlations were not significant, although were moderate with Assessor 1 (0.659 and 0.537).

### Conclusions

This was one task of six in a unit of work that was selected for inclusion in this research as a sample examination. Students did not view it as a separate task but just one activity in a series that culminated in their design and development of a prototype. In this context their attitude toward the theory/practical balance was relevant to the extent that they preferred to engage in practical work than in theory and planning. Apart from that the students responded well to the style of exam and generally preferred it to a theory paper-based exam.

The student development of this task as part of their portfolio provided adequate scope for a range of assessments. The students were allocated time in class to do their research, but some found the limited Internet access at the school frustrating.

Capturing the performance of the students was relatively straightforward, downloading each student’s folder from the school web system onto a CD and then uploading this into the MAPS portfolio system. Students could choose the format of their portfolio presentation, and they used Word and PowerPoint. Some students developed separate files for each aspect of the task, resulting in a submission of a number of Word and PowerPoint files in the folder. This rendered the marking exercise more complex and time consuming.

In terms of the marking of the student work the following conclusions could be drawn:

- Even using the same outcome descriptors, there was no correlation between the markers and the teacher awarding of levels to the students.
HE: Extended Production Exam

The HE case study involved one class of Year 11 students at one school completing the Engineering Studies course Units 2A-2B. There were 19 students across two classes that were in electronics Engineering context.

The aim was for each student to collate a portfolio of work, stored in digital formats, that was a part of a larger go-cart design and production task. The extended production examination ran for one semester, and the design portfolio that was examined as part of this project was one of the tasks within the task, together with a case study, research assignment, major project, workbook and examination.

The design portfolio documented the designing and making of the students design. The theme was an automated model wheelchair that uses digital inputs to control motion and direction. The students had to design and build a remote control for a buggy and build a prototype model. They did a case study in which they looked at wheelchairs and from there started making a prototype. Control circuitry used a PICAXE 28 microcontroller as its hub. Structural parts were milled from acrylic using an N/C milling machine and additional parts were designed/selected as required.

The portfolio included:

1) Problem statement
2) Systems requirements
3) PCB mask
4) PCB component overlay
5) PCB component list
6) Manufacture of PCB
7) Materials used by PCB
8) Structural parts
9) Basic control
10) Evaluation

The assessment addressed Outcomes 1 and 2 of the Engineering course, and used the related levels of achievement. The teacher assessed the work for the purposes of student grades. Two external assessors assessed the work as part of the research being undertaken. This assessment did not influence the assessment of the teacher but was compared and analysed. The teacher developed the tasks and portfolio template based on outcome and course requirements.

Students were provided structured time to enable them to progress through the range of activities in this task. The design portfolio, which was the one task assessed in this project, was developed over a number of weeks in a systems workshop which was equipped with computers. Students could concurrently work on the practical aspects of the task and the documentation of their activities through their portfolios. A portfolio template was not provided to the students but a content outline was given, together with some details, such as the PCB mask specifications and top and bottom plate design. All the portfolios were constructed using a Power point platform.

Results from data analysis

A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class

The researcher visited the classes to observe them working on their portfolio. They were allowed some variability in their use of class time, and were working in a systems lab, so could work on their models as well as on the computer. It was quite a free environment in which students could assist each other.
Survey of Students

There were 17 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.

Table 5.5

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>17</td>
<td>2.8</td>
<td>3.9</td>
<td>3.4</td>
<td>0.35</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>17</td>
<td>1.4</td>
<td>2.8</td>
<td>2.3</td>
<td>0.39</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>17</td>
<td>2.0</td>
<td>2.8</td>
<td>2.4</td>
<td>0.23</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confid</td>
<td>17</td>
<td>2.1</td>
<td>2.7</td>
<td>2.5</td>
<td>0.19</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>17</td>
<td>2.3</td>
<td>4.0</td>
<td>3.6</td>
<td>0.50</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>17</td>
<td>21</td>
<td>180</td>
<td>66.4</td>
<td>38.39</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

Some students had done examinations on computer before, but felt they needed little time to get used to the process. The felt comfortable using computers, and that the computer was appropriate for this task. They were least positive about the question item related to the computer enabling them to reflect on their ideas. They all have a range of IT related equipment at home and all the students have internet broadband. Most of the students use their home computers daily, and use the school computers for about an average of a little over an hour each day.

Most students enjoy using computers and felt positive about their role in school and society. Most felt quite confident about their skill levels in the range of skills listed in Q13. Students indicated a high level of ICT skills across the range of applications.

Figure 5.9: Distribution of scores for the Skills and SCuse scales on the student questionnaire.
Students Forum

The students interview forum was conducted by another researcher and analysed in terms of: what the students thought about the assessment; what they thought the attitude of other students was; whether the form of assessment allowed them to do their best work; what they would like changed in the future; any technical problems encountered in doing the task; and any other thoughts they had about doing assessment on computers.

The students believed that the unit of work was quite good. It took them about two terms but there were a few lessons that they spent on other things. They commented that they felt that other students would enjoy the work. They said that they enjoyed the fact that it was different, as many other previous test were just done on paper. They were enthusiastic about doing more digital work similar to this task.

In terms of technical difficulties, some students had problems with soldering on the board. However, the students aid that nobody had problems with the computers.

The students felt that this examination system was easier for a number of reasons and faster than writing or presenting work in other ways. They thought the examination period was too long and were concerned about losing data, but this seemed not to have occurred. Some recognised that they were easily distracted from the task at hand while working on the computer.

Teacher Structured Interview

A structured ‘interview’ was carried out via email with the teacher employing the task in the school. The questions forming the interview covered what they thought of the task and its structure and potential, how the students reacted to the activity and the quality of their output, and whether there were any technical or other problems with the implementation.

The teacher believed the task was a valuable tool to assess students’ level of understanding and familiarity, frequency of use, and level of skills. They felt that the task had clear instructions and the timing of the tasks was correct. The students readily accepted the task as part of the normal school routine while recognising that it was unusual and their work output was the usual standard with general to high success in the production of folios. The teacher commented that the quality of work produced was variable in the same way as the quality may vary with a traditional task. Better students produced clearer, well set-out material while others produced sketchy, less detailed output. In terms of potential, the teacher suggested that an increasing number of assessments should be carried out digitally across subjects.

Overall, the teacher believed that students enjoyed the more practical aspects of the assessment task while recognising that there was value in generating journal entries in a folio format to document their progress on a project.

The teacher did not report any technical problems with the activity. However, they emphasised that it is important that full details as a set of student instructions are clearly available at the beginning. Other than the possibility of student plagiarism there were no problems with implementing the task.

The teacher suggested the possibility of students being initiated into these types of tasks in middle to upper primary school. This should be done across the curriculum and take a variety of forms.
Results of Marking

For the standards-based marking method the two assessors and the teacher used the rubric provided in the course document based on Outcomes 1 and 2, as described below to mark the 19 students’ work. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the following instructions on the criteria to assess.

<table>
<thead>
<tr>
<th>Outcome/Aspect</th>
<th>Assessors’ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Engineering Process: Students apply and communicate a process to design,</td>
<td>Problem statement to investigate design needs and opportunities</td>
</tr>
<tr>
<td>make, and evaluate components and systems.</td>
<td>Systems requirements</td>
</tr>
<tr>
<td>Aspect 1: Students investigate design needs and opportunities in engineering</td>
<td>PCB mask, component overlay and component list for 'motherboard'</td>
</tr>
<tr>
<td>Aspect 2: Students generate engineering production proposals to solutions</td>
<td>PCB mask design and component list for L293D motorcontroller: Generate engineering production</td>
</tr>
<tr>
<td>Aspect 3: Students manage engineering production processes to produce solutions</td>
<td>Description of PCB manufacture process</td>
</tr>
<tr>
<td>Aspect 4: Students evaluate intentions, plans and actions</td>
<td>Evaluation of intentions, plans and actions</td>
</tr>
<tr>
<td>2 - Engineering understandings: Students understand properties of materials,</td>
<td>PICAXE 28 PCB: Properties of materials and/or components</td>
</tr>
<tr>
<td>energy transfer and design principles in engineering technologies.</td>
<td></td>
</tr>
<tr>
<td>Aspect 1: Students understand properties of materials and/or components in</td>
<td></td>
</tr>
<tr>
<td>engineering technologies</td>
<td></td>
</tr>
</tbody>
</table>

The two external assessors and the teacher marked the student work on the aspects of the two outcomes (1 & 2). All the judgements on the two outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12. The numeric scores for each assessor were summed across the outcomes and averaged. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

<table>
<thead>
<tr>
<th>Student</th>
<th>Assessor A1</th>
<th>Assessor A2</th>
<th>Average (Avg)</th>
<th>Teacher (Tch)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE101</td>
<td>17.0</td>
<td>17.5</td>
<td>17.25</td>
<td>14.0</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td>5.0</td>
</tr>
<tr>
<td>HE102</td>
<td>18.0</td>
<td>17.5</td>
<td>17.75</td>
<td>16.0</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td>4.25</td>
</tr>
<tr>
<td>etc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.0</td>
</tr>
</tbody>
</table>

Figure 5.10: Tabulation of results from marking.
A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

<table>
<thead>
<tr>
<th>Correlations of Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N=10)</td>
</tr>
<tr>
<td>TOT_M1</td>
</tr>
<tr>
<td>Pearson Correlation</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
</tr>
<tr>
<td>TOT_M2</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
</tr>
<tr>
<td>Tch</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

<table>
<thead>
<tr>
<th>Correlations of Rankings</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N=10)</td>
</tr>
<tr>
<td>Rank of TOT_M1</td>
</tr>
<tr>
<td>Pearson Correlation</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
</tr>
<tr>
<td>Rank of TOT_M2</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
</tr>
<tr>
<td>Rank of Tch</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Two external markers used the aspects to make level statements for the two outcomes to mark the exam outputs. These markers had a significant level of agreement with an initial inter-rater reliability correlation of 0.786 (p<0.01) on the marks of students, and 0.766 (p<0.01) on the ranks of students.

The teacher marked the students work using outcomes and levels. These were then converted to a mark between 12 and 25 to represent levels between 4 and 8. Comparison with external marker results indicated a significant agreement correlation of 0.457 (p<0.05) on the marks and 0.630 (p<0.01) on the ranks of the students.

**Conclusions**

Overall it was concluded that this extended production exam was implemented with no significant technical difficulties with all students able to complete the requirements of the exam over the time permitted. The students were engaged in a range of concurrent tasks over this period, so their time was not singularly devoted to the portfolio that was examined in this project.

The students responded well to the style of exam and generally preferred it to a theory paper-based exam although they mainly liked the ease of use in using a computer, but thought it went on for too long (1 term). Capturing the performance of the students was relatively straightforward as the students were saving their files to the school server and then upon completion they were downloaded to a DVD. The provision of a portfolio template complete with key data may have limited the range of capabilities reflected in the portfolio.

In terms of the marking of the student work the following conclusions could be drawn:

- There was a higher correlation between the assessors than with the teacher, although this was still significant.
- There was a high level of agreement between the two external assessors using the aspects descriptors.
YE: Focussed performance task

The YE case study involved one class of Year 11 students at one school completing the Engineering Studies course Unit 2A. There were 7 students in the class that was in a gears design context.

The aim was for each student was to create two meshing gears according to specific criteria, which involved both researching, designing and making. This activity was one part of a more extensive unit on the materials, structures and mechanical systems option in Engineering Studies. The students were required to submit all their work to the Submit folder on the school server as a digital portfolio in Powerpoint format. The assessment would address Outcomes 2 and 3 of the Engineering course and use the related levels of achievement. The teacher initially used the grade descriptors to allocate marks of Satisfactory, High and Very High to the students’ performance, then re-marked allocating levels to enable comparisons with the external assessors. The teacher assessed the work for the purposes of student grades. Two external assessors also assessed the work as part of the research being undertaken. This assessment did not influence the assessment of the teacher but was compared and analysed.

The teacher developed the task based on materials provided by the Curriculum Council. The instructions to students were to:

- Using Prodesktop, Modela and the CNC machine create two acrylic meshing gears that will produce a 50% reduction in speed.
- The students were instructed to follow a set procedure as follows:
- Design the gear teeth and calculate the diameter of the gear wheels.
- Draw the gear wheels on ProDesktop and create an assembly to test the meshing of the gears.
- Produce an A4 engineering drawing of the two gears separate.
- Save in a Modela friendly version.
- Prepare the cutting codes for the CNC machine in Modela.
- Set up CNC machine and initiate cutting.
- Make a gear mounting block using pine and 6mm dowel to demonstrate the efficiency.
- Present a portfolio consisting of all of your work to the ‘submit folder’ using PowerPoint.

The students were told that they would need to take digital photographs as they progress through the practical work and to record their work with the CNC machine in a digital video format. The content of their submission was to include:

- Gear wheel design work (may need to be scanned)
- Calculations
- Gear wheel drawing
- Test assembly drawing
- Engineering drawing
- Video of CNC process (set up and start)
- Picture of gear mounting block
- Evaluation

Students were provided with a number of weeks of class time for this task. Many students were not able to complete the task for a number of reasons. The main reason was the limitation for machining the gear wheels as there was only one CNC machine available for this aspect of the task and some students required considerable time in setting up and testing and then running their designs.

The limited availability of video and still digital cameras also prevented students from documenting the full extent of their production. Consequently many of the students’ submissions were incomplete in this regard.
Results from data analysis
A range of data was collected and analysed, including observation of the classes, an interview with a group of students, an interview with the teacher, a survey of the students, and the output from their portfolio.

Observations of the class
The researcher visited the class to observe the students working on the task. They were quite independently working on a range of tasks, managing their own time in working toward completion. Some were assisting each other and others worked independently.

Survey of Students
There were 15 students who completed the survey. The questionnaire consisted of 58 closed response items and two open-response items. The minimum, maximum, mean and standard deviation were calculated for each closed response item using SPSS. Responses to the open-response items were tabulated to assist in drawing out themes. A number of scales were derived from combining items from the questionnaire. Results are shown in the table and graphs below.

Table 5.6
Descriptions and descriptive statistics for the scales based on items from the student questionnaire (N=15).

<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. D</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAssess</td>
<td>15</td>
<td>2.4</td>
<td>3.6</td>
<td>3.2</td>
<td>.28</td>
<td>Ease of assessment for the students (combination of all items in Question 2). Score between 1 and 4.</td>
</tr>
<tr>
<td>Apply</td>
<td>15</td>
<td>1.6</td>
<td>3.0</td>
<td>2.3</td>
<td>.46</td>
<td>Application of some computer uses (combination of all items in Question 10). Score between 1 and 4.</td>
</tr>
<tr>
<td>Attitude</td>
<td>15</td>
<td>1.8</td>
<td>3.0</td>
<td>2.4</td>
<td>.30</td>
<td>Attitude towards using computers (combination of all items in Question 11). Score between 1 and 4.</td>
</tr>
<tr>
<td>Confid</td>
<td>15</td>
<td>1.5</td>
<td>3.0</td>
<td>2.4</td>
<td>.40</td>
<td>Confidence in using computers (combination of all items in Question 12). Score between 1 and 4.</td>
</tr>
<tr>
<td>Skills</td>
<td>15</td>
<td>2.4</td>
<td>4.0</td>
<td>3.2</td>
<td>.52</td>
<td>A measure of ICT skills (combination of all items in Question 13). Score between 1 and 4.</td>
</tr>
<tr>
<td>SCuse</td>
<td>15</td>
<td>28</td>
<td>120</td>
<td>73.4</td>
<td>33.42</td>
<td>Estimate of time per day spent using computers at school (combination of all items in Question 8).</td>
</tr>
</tbody>
</table>

Students had not often done a test using a computer before but felt they would only need a little more time to get used to it. They generally felt it was easy to use IT in their project but were less sure that
IT was good for reflecting on their task. Most students had the basic IT equipment at home, all had Internet and most had broadband. They used their home computer most days and spent an average of 1-1½ hours each day on computers at school. They felt quite positive about using computers for schoolwork and generally about IT. They felt quite confident about the skills listed in Q13, though less confident about databases and web authoring.

**Students Forum**
The students described the task in the following way:

“We had to make two cogs that were a certain ratio and design them using Pro-Desktop. We had a specific amount of plastic that we could integrate into the manufacturing of the cogs. We had to use the CNC machine to cut out the cogs.”

They believed the task was a good exercise, better than doing a test/exam and most students had a positive experience. The students said that they felt confident that they could complete the task in the allocated time. They preferred this task to their previous assessment that had been a test on paper.

The students appreciated that the work was more practical and hands on and they felt that by making the gear straight after the designing stage they would be more familiar with what they had done. They reported that there were no problems during the design stage. Each student did the task individually when using the CNC machine. They were not allowed to watch others doing the assessment.

**Teacher Structured Interview**
A structured ‘interview’ was carried out via email with the teacher employing the task in the school. The questions forming the interview covered what they thought of the task and its structure and potential, how the students reacted to the activity and the quality of their output, and whether there were any technical or other problems with the implementation.

The teacher believed the task was valuable and effective and that it linked well with current class activities and used an appropriate level of technology. The main part of the assessment activity was completed in one school period. However, the practical section took a lot of time. Therefore the teacher felt that they would need to start earlier in the year if it is to be effective. They also commented that the students reacted very well to the activity and were interested and positive producing satisfactory to good quality work given the timeframe available to them. The student performance was not surprising to the teacher. They were please with some students’ work but disappointed with others.

The teacher believed the activity could work at their school. However they were concerned that the full activity required a computer numerically controlled machine. They were confident that this type of assessment has potential for other subject areas.

The teacher made the point that the students preferred the task to a three-hour theory examination. The only problem was that occurred was that the class only had access to one computer numerically controlled machine and thus the students had to take turns meaning the activity ran over several weeks. There were no other problems implementing the activity. In terms of suggestions for developing the use of digital forms of assessment the teacher reported that students were creating their own websites and posting work on them. They then send the teacher the URL so the work can be browsed and assessed.

**Results of Marking**
For the standards-based marking method the two assessors and the teacher used the rubric provided in the course document based on Outcomes 2 and 3, as described below top mark the 13 students’ work. For the overall outcome and each ‘Aspect’ of the outcome levels of achievement are described for Levels 4 through to 8. The assessors were given the following instructions on the criteria to assess.
Outcome/Aspect | Assessor's Criteria
---|---
2. Engineering Understandings: Students understand properties of materials, energy transfer and design principles in engineering technologies.
Aspects:
- Investigate design needs and opportunities in engineering
- Generate engineering production proposals to solutions
The student:
- generates a range of possible solutions;
- uses appropriate graphical and technical language.

3. Engineering technology skills: Students use materials, skills and technologies appropriate to the engineering industry.
Aspects:
- Apply initiative and organisational skills
- Apply materials, techniques, and technologies to achieve solutions to engineering challenges
The student:
- works to specifications;
- uses an appropriate level of technical language;
- manages time to ensure completion of the task.

The two external assessors marked the student work on the aspects of the two outcomes (2 & 3) while the teacher also marked using levels. All the judgements on the two outcomes were converted to numbers (ranging from 12 to 25 per outcome) using the Curriculum Council’s converter with the lowest Level 4 equating with 12. The numeric scores for each assessor were averaged across the outcomes. For each assessor and for the teacher the students were ranked so that comparisons could be made. The results for each student were tabulated using the headings shown in the figure below.

![Figure 5.10: Tabulation of results from marking.](image)

A correlation analysis was done on the rankings and on the scores generated. The results are shown in the figure below.

### Correlations of Marks

<table>
<thead>
<tr>
<th>(N=15)</th>
<th>TOT_M1</th>
<th>TOT_M2</th>
<th>AVG</th>
<th>Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT_M1 Pearson Correlation</td>
<td>1</td>
<td>.979(**)</td>
<td>0.9895</td>
<td>.119</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
<td></td>
<td>.800</td>
</tr>
<tr>
<td>TOT_M2 Pearson Correlation</td>
<td>.979(**)</td>
<td>1</td>
<td>0.9895</td>
<td>.135</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
<td></td>
<td>.774</td>
</tr>
<tr>
<td>Tch Pearson Correlation</td>
<td>.119</td>
<td>.135</td>
<td>0.127</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.800</td>
<td>.774</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* **Correlation is significant at the 0.01 level (2-tailed).*

### Correlations of Rankings

<table>
<thead>
<tr>
<th>(N=15)</th>
<th>Rank of TOT_M1</th>
<th>Rank of TOT_M2</th>
<th>Rank AVG</th>
<th>Rank of Tch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank of TOT_M1 Pearson Correlation</td>
<td>1</td>
<td>1.000(**)</td>
<td>1</td>
<td>.158</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
<td></td>
<td>.735</td>
</tr>
<tr>
<td>Rank of TOT_M2 Pearson Correlation</td>
<td>1.000(**)</td>
<td>1</td>
<td>1</td>
<td>.158</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
<td></td>
<td>.735</td>
</tr>
<tr>
<td>Rank of Tch Pearson Correlation</td>
<td>.158</td>
<td>.158</td>
<td>0.158</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.735</td>
<td>.735</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* **Correlation is significant at the 0.01 level (2-tailed).*

*Figure 5.11: Correlations between results from marking YE student work.*
Two external markers used the level statements for the two outcomes to mark the student work. These markers had a strong level of agreement with an initial inter-rater reliability correlation of 0.979 (p<0.0001) on the marks given for the students’ work, and an exact match when these marks were converted to a rank.

The teacher marked the students work using outcomes and levels. These were then converted to a mark between 12 and 25 to represent levels between 4 and 8. Comparison with external markers results indicated little agreement, reflected in a low and non-significant correlation for both the mark and the rank. The assessors’ marks were a little higher than the teachers, ranging from 12 – 14.5, while the teacher’s marks ranged from 11 – 13.

Conclusions
Overall there were some significant difficulties with this focussed performance task. The dependence on limited equipment (in this case a CNC machine) essential to the task was an impediment to all students’ successful completion.

The range of file formats students were using was also problematic. Each student submission arrived in a folder with up to five separate files all of different formats. Some were scanned documents and a number were screen captures of various aspects of the machining processes so were in a format unique to the machine software. If the students had more time this difficulty could have been overcome, as it was the researchers placed all the files in a PDF document in order to provide an overall picture of each students capability.

Nevertheless, the students responded well to the focussed performance task as a part of their assessment and appreciated the fact that the examination was measuring their actual performance rather than their recall or knowledge.

The predictable nature of the student outcomes because of limited design opportunities, and the consequent fact that their submissions were all very similar, provided little scope for a broad range of marks, and also few opportunities for students to demonstrate a high level of achievement. Some student submissions included very little personalised annotation, just a series of screen captures and technical drawings and little evaluation. It is concluded that a lack of time was the most problematic factor in limiting student achievement.

In terms of the marking of the student work the following conclusions could be drawn:

- There were clear differences in results between the two external assessors and the teacher, despite using the same assessment guidelines.
- There was a high level of agreement between the two external assessors.
- No students were judged by any of the assessors to have demonstrated any work above L4.
6 - Findings, Conclusions and Recommendations

A range of types of quantitative and qualitative data was collected at each school contributing to each case study. These data were analysed, to address the research questions, within a feasibility framework of the four dimensions described below.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Description</th>
<th>Types of Data Collected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manageability</td>
<td>Concerning making a digital form of assessment do-able in typical classrooms with the normal range of students.</td>
<td>Observation, student surveys, student interviews, teacher interviews and reports.</td>
</tr>
<tr>
<td>Technical</td>
<td>Concerning the extent to which existing technologies can be adapted for assessment purposes within course requirements.</td>
<td>Observation, teacher interviews and deliberations within research team</td>
</tr>
<tr>
<td>Functional</td>
<td>Concerning reliability and validity, and the comparability of data with other forms of assessment.</td>
<td>Interviews with teachers and assessors, quality of digital representations.</td>
</tr>
<tr>
<td>Pedagogic</td>
<td>Concerning the extent to which the use of a digital assessment forms can support and enrich the learning experience of students.</td>
<td>Questionnaires and interviews with teachers and students.</td>
</tr>
</tbody>
</table>

In addition conclusions were drawn with respect to the constraints and benefits of the form of assessment used in each case study.
### 6.1 Conclusions from AIT Case Studies

Conclusions from the AIT case studies are based on the analysis provided in the table on the following two pages. A summary of findings from the combination of cases is presented in the table below.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manageability</td>
<td>There were three case studies involving ‘practical exams’ with two (1 and 3 hour exams) providing no technical or logistical difficulties and in one there were students who lost work and there were logistical difficulties in managing student time over a two-week period. Of the four case studies involving portfolios in only one case were the portfolios relatively easy to access and mark. This case involved all students using a standard web-based template. The other cases were difficult either due to the range of software used, the quantity of items required, a lack of time for completion, a lack of supporting documentation, or in one case a lack of student experience and ICT capability.</td>
</tr>
<tr>
<td>Technical</td>
<td>Student work was collected either on a class DVD or USB drives and submitted either to a server or directly to the USB drives. Some degree of maintenance was required for all four cases involving portfolios with the least for the template based web format. The three cases of practical exams typically only required conversion to PDF. All work in PDF, web or Flash format was easily accessed by markers.</td>
</tr>
<tr>
<td>Functional</td>
<td>The two external assessors used a marking rubric developed from a standards-based framework. This resulted in highly correlated marks and rankings in all cases. The comparative pairs method of marking was successfully implemented for a short performance exam with resulting highly reliable scores but questions remain concerning the validity of results with variations for particular students when cut scores were applied compared with the standards-based assessment. In only two cases (2 week production exam and low quality portfolios) were the ranking of students highly correlated between the teacher and assessors, in one case despite differences in ‘marks’. In all cases there were significant differences in the marks produced by the teacher and the assessors. In three cases this was due to the use of different outcomes and in two cases due to the use of analytic numerical marking method. In two cases differences were likely to be due to the teacher having access to more. In all cases inadequate documentation and annotation by students limited achievement to L5 or below. Further this led to differences between external assessors due to interpretation of student intention. In most cases increased scaffolding and control of time use may have partly addressed this issue. In two of the ‘exams’ and two of the portfolios the structure of the tasks limited the opportunity to demonstrate above L4. In five of the cases a lack of time limited the opportunity to demonstrate above L4.</td>
</tr>
<tr>
<td>Pedagogic</td>
<td>Typically students liked doing practical work and disliked documenting work or answering question. In at least three cases students liked learning ICT skills through the assessment. Most students demonstrated L4 work, in some cases many did not, and in some cases a few demonstrated L5 work. In general the quality of work was low probably largely as a result of the capability of the students.</td>
</tr>
</tbody>
</table>
6.1.1 Conclusions About Exams

Two of the exams were production exams and one was a performance task exam. The performance task exam was short, easy to manage and mark but was limited in scope for assessment of student capability. The longer production exams provided scope for assessment but lacked sufficient scaffolding to ensure students provided adequate explanation of their intentions in design and production. The degree of structure was an issue where too little structure risked students not including enough to assess and too much limiting the opportunity to demonstrate higher-level achievement. In all three cases greater control over student time use would have increased the quality of evidence of achievement. That is, students needed a clear allocation of time for annotation, documentation, and production. Another issue was the level of access students were permitted to information and resources. In two cases students had unlimited access that in one case distracted them from completing the requirements while in the other case was only used to access graphic images. In the third case students only had access to software help files and their exam folders and that probably reduced the effectiveness of their designs.

The three cases investigated were:

1. Production of a poster (50 minutes)
2. Production of a web-page with responses to some questions (3 hours)
3. Production of a video with technology process documentation (2 weeks)

Conclusions to date:

- All three were implemented successfully with the poster production the least problematic. In only one case did a computer crash (video production) with the student losing some work.
- There is a fine balance between providing adequate structured scaffolding and providing adequate flexibility to allow demonstration of achievement. The Poster exam had very little scaffolding and thus marking required some ‘guesswork’ on student intention. The video exam had a lot of scaffolding that constrained students to the extent that none of them entirely finished.
- Marking by ‘adding up marks for items’ results in significantly different rankings than using standards-based rubrics. The latter provide reliable results that were well correlated with comparative pairs marking rankings.

6.1.2 Conclusions About Portfolios

The only set of portfolios that was relatively easy to mark was template based and in a web format. Where students were given considerable choice in either format/software or structure/style marking became more difficult and unreliable. The teacher’s knowledge of the student and other work completed by the student was likely to influence results compared with those from the assessors. It was critical that students did not have to spend a lot of time learning to use the construction software, however, this limited the pedagogic value. The degree of structure was also an issue where too little structure risked students not including enough to assess and too much limiting the opportunity to demonstrate higher-level achievement. However, it appeared to be better to err on the side of too much structure.

The four cases investigated were:

1. Template-based web-site including 4 outcomes and 4 tasks
2. Tasks-based and coordinated through Powerpoint (SIDE module)
3. Skills-based in Flash with links to other tasks
4. Outcomes-based in student chosen format (e.g. Word, Powerpoint, Flash, website)

Conclusions to date:

- Template-based website is easiest to collect and to mark. All files in one folder, including an index page, then zipped and uploaded into MAPS. As well as providing an interface and means of navigation it is likely that a multi-page structure is preferable to a tagged single page.
Portfolios in Flash are easy to manage when published for the web. Once again all linked files need to be zipped within one folder.

Tasks-based portfolio was difficult to manage and mark, particularly as no student had successfully linked all the required items to a slideshow.

Insisting on a particular structure given in a template with requirement for annotations is easiest to mark. Few students provided adequate annotations for any work and would need more prompting.

Limiting file formats and the size of media files is essential. For example, where possible convert files to HTML, PDF, JPG, GIF or AVI. If a slideshow was provided in PDF format then in the annotation the student could draw attention to a menu structure (may not be necessary to actually try the slideshow). Alternatively an animated slide could be recorded as a WAV file.

Marking using standards-based rubrics provides reliable rankings.

### 6.1.3 Conclusions About Methods of Marking

The standards-based method of marking using a rubric was successfully implemented with the support of a digital marking tool and an online portfolio system (MAPS) with the marks from the two assessors, and resulting rankings, highly correlated in all cases but rarely highly correlated to the teacher marks. For all cases there were only a few students whose work was judged differently by the two assessors and required a consensus meeting to resolve.

<table>
<thead>
<tr>
<th>Case</th>
<th>Type of Assessment</th>
<th>Correlations Between Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Assessors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marks</td>
</tr>
<tr>
<td>WA</td>
<td>Digital portfolio: web-based template</td>
<td>0.87**</td>
</tr>
<tr>
<td>AA</td>
<td>Digital portfolio: slideshow</td>
<td>0.87**</td>
</tr>
<tr>
<td>IA</td>
<td>Digital portfolio: Flash movie</td>
<td>0.97**</td>
</tr>
<tr>
<td>PA</td>
<td>Digital portfolio: choice of format</td>
<td>0.89**</td>
</tr>
<tr>
<td>MA</td>
<td>Digital video production exam (2 weeks)</td>
<td>0.83**</td>
</tr>
<tr>
<td>CA</td>
<td>Web page production exam (3 hours)</td>
<td>0.90**</td>
</tr>
<tr>
<td>XA</td>
<td>Poster graphics task exam (50 mins)</td>
<td>0.99**</td>
</tr>
</tbody>
</table>

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

The comparative pairs method of marking was successfully implemented for a simple short production exam with resulting highly reliable scores but questions remain concerning the validity of results with variations for particular students compared with the standards-based assessment. Assessors took between 2 and 3 hours to make the comparisons for 48 scripts although it was determined that the number of comparisons may have been able to be reduced by half. It is likely that the validity may have been improved through increased training for the assessors and through providing a note-taking facility with the digital marking tool.

### 6.2 Summary of Findings from AIT Case Studies

A summary of findings from the AIT case studies was compiled based on the Feasibility Framework and including a summary of the constraints and benefits of the form of assessment used in the case study. The Functional dimension was divided into findings regarding Validity and those regarding Reliability. Validity was analysed by considering: (1) how well the performance of students matches the curriculum outcomes; (2) the extent to which the method of representing performance was authentic; and, (3) whether the task and context were meaningful and relevant to students and community practice.
6.3 Conclusions from Engineering Case Studies

Conclusions from the Engineering Studies case studies are based on the analysis provided in the table on the following two pages. A summary of findings from the combination of cases is presented below.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manageability</td>
<td>There were three case studies involving ‘practical exams’ with all providing no technical or serious logistical difficulties. In one case there were some difficulties in managing student time on computers and in two cases there were difficulties accessing student work due to the range of file formats and incompatibilities between file types. Of the three case studies involving reflective portfolios stored on a server in a structured manner making them easy to access. In one case students did not have enough time.</td>
</tr>
<tr>
<td>Technical</td>
<td>In all but one case student work was collected on a server and later burnt on an optical disc. In one case USB drives were used. In three cases incompatibility between file types and/or size of files were a difficulty.</td>
</tr>
<tr>
<td>Functional</td>
<td>In all but one case students readily perceived the assessment task to be authentic and meaningful. In four cases the task was structured permitting a good range of levels of achievement to be demonstrated. Where the two external assessors used a marking rubric developed for the task from a standards-based framework this generally resulted in correlated marks and rankings. In no cases were the marks or ranking of students highly correlated between the teacher and assessors. In four cases there were significant differences in the marks produced by the teacher and the assessors. This was due to different interpretations of the outcomes. In two cases the differences were likely to be due to the teacher having access to more information about the student performance.</td>
</tr>
<tr>
<td>Pedagogic</td>
<td>Typically students liked doing practical work and disliked documenting work or answering questions. Assessment tasks worked best where the approach was familiar to the students with in one case this being part of ‘the culture’. Most students demonstrated at least L4 work, and only in a few cases less than L4</td>
</tr>
</tbody>
</table>

6.3.1 Conclusions About Exams

There were three cases involving production exams. They ranged in time from four hours to one semester. In each case the evidence of performance from the exam was gathered into a digital portfolio.

The longer production exams provided scope for assessment but in some instances lacked adequate scaffolding to ensure students provided an explanation of their intentions in design and production. The degree of structure was an issue where too little structure risked students not including enough to assess and too much limiting the opportunity to demonstrate higher-level achievement and a diversity of achievement. In all three cases greater control over student time use would have increased the quality of evidence of achievement. That is, students needed a clear allocation of time to work on their production, documentation and research.

The three cases investigated were:

1. Design and production of lock mechanism for a pool gate (4 hours)
2. Design and production of a go-kart (1 semester)
3. Design and production of two gear wheels (3 weeks)

General conclusions were:

- All three were implemented successfully with the go-kart the least problematic.
- In two of the cases there were difficulties in accessing the students’ work due to the range of file formats used and incompatibilities between file formats and software.
• For one case there were not enough computers for students to create their evidence portfolios.
• In all cases the assessment task was seen to be authentic (i.e. a relevant context for students, real tasks done in real time).
• In two cases task rubrics were developed for marking that resulted in highly correlated results for the two assessors but not with the teacher. In the other case only the course outcome scales of achievement were used and there was no significant correlation in marks.
• In one case this type of assessment was ‘part of the culture’ and was the best accepted by the students with the least difficulties and well correlated results. The other two cases required the teachers to provide a lot of instruction and student results varied according to the level of student efficiency and organisation.
• There is a fine balance between providing adequate structured scaffolding and providing sufficient flexibility to allow demonstration of achievement. The go-kart exam had very little scaffolding and thus marking required some ‘guesswork’ on student intention. The pool gate exam had a lot of scaffolding that enabled all students to end up with a representative product.

6.3.2 Conclusions About Reflective Portfolios

There were three cases involving reflective portfolios.

The portfolios that were most easy to mark were template based. Where students were given considerable choice in either portfolio format or structure/style, marking became less consistent and unreliable. The teacher’s knowledge of the student and other work completed by the student was likely to influence results compared with those from the assessors. The degree of structure was also an issue where too little structure risked students not including enough to assess and too much limiting the opportunity to demonstrate higher-level achievement. However, it appeared to be better to err on the side of too much structure.

The three cases investigated were:

1. Template-based Powerpoint portfolio of research for an automated wheelchair (1 semester)
2. Template-based Powerpoint portfolio of research for an electronic toy (1 semester)
3. Powerpoint portfolio of research related to Engineering mechanisms (1 semester)

General conclusions were:

• There were no technical difficulties and the only logistical difficulty was a lack of time for the template-based slideshow
• All student work was saved on the school’s server and later transferred to an optical disc. In one case there were file size issues for the server.
• In all cases there was little, if any, correlation between the two assessors marks and none between them and the teacher.
• It was important that students were able to connect the value of the portfolio to later practical work. In one case this did not occur with a lot of written work that students could not see directly relating to practical outcomes.
• In all cases a lack of time was a serious constraint on the value of the assessment and the opportunity for students to represent their capability. However, in two cases the students were still positive about this approach to assessment and definitely preferred it to a written exam.
• Powerpoint portfolios were easiest to collect and mark.
• Tasks-based folders, with work submitted in a number of formats were difficult to manage and mark, particularly as no student had successfully linked all the required items to a slideshow.
• Insisting on a particular structure given in a template is easiest to mark.
6.3.3 Conclusions About Methods of Marking

The standards-based method of marking using a rubric was successfully implemented with the support of a digital marking tool and an online portfolio system (MAPS) with the marks from the two assessors, and resulting rankings, highly correlated in three cases but never highly correlated to the teacher marks. There was little negotiation between assessors, and between assessors and the teacher on student work exemplars which represented the level statements, resulting in more diversity in assessment judgements than might have otherwise been the case. In those instances where common indicators of the levels of student work were utilised, correlations between the assessors was greater.

Table 6.2
Correlations between assessors and between assessors and teachers for the Engineering case studies.

<table>
<thead>
<tr>
<th>Case</th>
<th>Type of Assessment</th>
<th>Assessor criteria</th>
<th>Teacher marking</th>
<th>Asessors</th>
<th>Consensus and Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Marks</td>
<td>Rankings</td>
<td>Marks</td>
</tr>
<tr>
<td>GE</td>
<td>Digital portfolio: slideshow template</td>
<td>Rubric specific to task</td>
<td>Different to assessors</td>
<td>0.54*</td>
<td>0.45</td>
</tr>
<tr>
<td>ME</td>
<td>Digital portfolio: template</td>
<td>Rubric specific to task</td>
<td>Same as assessors</td>
<td>0.99**</td>
<td>1.00**</td>
</tr>
<tr>
<td>VE</td>
<td>Production Exam (2 x 2 hours): Pool gate</td>
<td>Course scales of achievement</td>
<td>Different to assessors</td>
<td>-0.36</td>
<td>-0.40</td>
</tr>
<tr>
<td>WE</td>
<td>Digital portfolio: student choice of format</td>
<td>Rubric specific to task</td>
<td>Different to assessors</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>HE</td>
<td>Production exam (1 semester): Go-carts</td>
<td>Rubric specific to task</td>
<td>Same as assessors</td>
<td>0.79**</td>
<td>0.77**</td>
</tr>
<tr>
<td>YE</td>
<td>Production Exam (3 weeks): Gears</td>
<td>Course scales of achievement</td>
<td>Different to assessors</td>
<td>0.98**</td>
<td>1.00**</td>
</tr>
</tbody>
</table>

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

6.4 Summary of Findings from Engineering Studies Case Studies

A summary of findings from the Engineering Studies case studies was compiled based on the Feasibility Framework and including a summary of the constraints and benefits of the form of assessment used in the case study. The Functional dimension was divided into findings regarding Validity and those regarding Reliability. Validity was analysed by considering: (1) how well the performance of students matches the curriculum outcomes; (2) the extent to which the method of representing performance was authentic; and, (3) whether the task and context were meaningful and relevant to students and community practice.
<table>
<thead>
<tr>
<th>Elements</th>
<th>Benefits</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course exam and assignment</td>
<td>Increased student interest and engagement</td>
<td>Time-consuming for instructors</td>
</tr>
<tr>
<td>Tutorials</td>
<td>Varied perspectives and peer learning opportunities</td>
<td>Limited student participation</td>
</tr>
<tr>
<td>Online quizzes</td>
<td>Immediate feedback and self-paced learning</td>
<td>Technology issues and access problems</td>
</tr>
</tbody>
</table>

**Summary of Findings from Engineering Case Studies**

- **NB:** P=Predictor, E=Exam, 1/2=2/12 Year 1/2 Students, N=Number of Students
- **TE** = Teaching Evaluation
- **EC** = Engineering Case Study
6.5 Overall Conclusions
In this section firstly the findings from the AIT case studies and Engineering case studies are summarised separately and then combined.

6.5.1 Summary of AIT findings
Manageability

• There were three case studies of ‘practical exams’ with two (1 and 3 hour exams) providing no technical or logistical difficulties and in one two students lost work and there were logistical difficulties in managing a two-week period.

• There were four case studies of portfolios with only one relatively easy to access and mark; it used a standard web-based template. The others were difficult to collect, access and mark either due to the range of software used, the quantity of items required, a lack of time for completion, a lack of supporting documentation, or a lack of student experience and ICT capability.

Technical

• Student work was collected either on a class DVD or USB drives and submitted either to a server or directly to the USB drives. There were no technical difficulties in this.

• Some degree of maintenance was required for all four cases involving portfolios with the least for the template based web format. Missing files and/or broken links were common.

• The three cases of practical exams typically only required conversion to PDF.

• Work in PDF, web or Flash format was easily accessed by markers.

Functional

• The two external assessors used a marking rubric developed from a standards-based framework. This resulted in highly correlated marks and rankings in all cases.

• The comparative pairs method of marking was successfully implemented for a short performance exam with resulting highly reliable scores but questions remain concerning the validity of results for particular students with variations when cut scores were applied compared with the standards-based assessment.

• In only two cases (2 week production exam and low quality portfolios) were the ranking of students by the teacher highly correlated with that by the assessors; in one case this was despite differences in ‘marks’.

• In all cases there were significant differences in the marks produced by the teacher and the assessors with little correlation for all but two cases. In three cases this was due to the use of different outcomes and in two cases due to the use of analytic numerical marking method. In two cases differences were likely to be due to the teacher having access to more.

• In all cases inadequate documentation and annotation by students limited achievement to Level 5 or below. Further, this led to differences between external assessors due to interpretation of student intention. In most cases increased scaffolding and control of time use may have partly addressed this issue.

• In two of the ‘exams’ and two of the portfolios the structure of the tasks limited the opportunity to demonstrate above L4.

• In five of the cases a lack of time limited the opportunity to demonstrate above L4.

Pedagogic

• Typically students liked doing practical work and disliked documenting work or answering question.

• In at least three cases students liked learning ICT skills through the assessment.
Most students demonstrated Level 4 work, in some cases many did not, and in some cases a few demonstrated Level 5 work. In general the quality of work was low probably largely as a result of the capability of the students.

Exams
- All three were implemented successfully with the poster production the least problematic. In only one case did a computer crash (video production) with the student losing some work.
- There is a fine balance between providing adequate scaffolding and providing sufficient flexibility to allow demonstration of achievement. The Poster exam had very little scaffolding and thus marking required some ‘guesswork’ on intention. The video exam had a lot of scaffolding that constrained students to the extent that none of them entirely finished.
- Marking by ‘adding up marks’ resulted in significantly different rankings than using standards-based rubrics. The latter provided reliable results well correlated with comparative pairs marking.

Digital Portfolios
- The template-based website is easiest to collect and to mark. All files in one folder, including an index page, then zipped and uploaded into MAPS. As well as providing an interface and means of navigation it is likely that a multi-page structure is preferable to a tagged single page.
- Portfolios in Flash were easy to manage when published for the web. All linked files needed to be within one folder before being zipped.
- The tasks-based portfolios were difficult to manage and mark, particularly as no student had successfully linked all the required items to a slideshow.
- Few students provided adequate annotations for any work and would need more prompting. A template with a structure requiring annotations was easiest to mark.
- Limiting file formats and size of media files is essential. Where possible files should be converted to HTML, PDF, JPG, GIF or AVI. If a slideshow was provided in PDF format then in the annotation the student could draw attention to a menu structure (may not be necessary for the assessor to actually try the slideshow). Alternatively an animated slide could be recorded as a WAV file.
- Marking using standards-based rubrics provides reliable rankings.

6.5.2 Summary of Engineering Findings

Manageability
- There were three case studies involving ‘practical exams’ with all providing no technical or serious logistical difficulties. In one case there were some difficulties in managing student time on computers and in two cases there were difficulties accessing student work due to the range of file formats and incompatibilities between file types.
- Of the three case studies involving reflective portfolios stored on a server in a structured manner making them easy to access. In one case students did not have enough time.

Technical
- In all but one case student work was collected on a server and later burnt on an optical disc. In one case USB drives were used.
- In three cases incompatibility between file types and/or size of files were a difficulty.

Functional
- In all but one case students readily perceived the assessment task to be authentic and meaningful.
• In four cases the task was structured permitting a good range of levels of achievement to be demonstrated.
• Where the two external assessors used a marking rubric developed for the task from a standards-based framework this generally resulted in correlated marks and rankings.
• In no cases were the marks or ranking of students highly correlated between the teacher and assessors.
• In four cases there were significant differences in the marks produced by the teacher and the assessors. This was due to different interpretations of the outcomes. In two cases the differences were likely to be due to the teacher having access to more information about the student performance.

Pedagogic
• Typically students liked doing practical work and disliked documenting work or answering questions.
• Assessment tasks worked best where the approach was familiar to the students with in one case this being part of ‘the culture’.
• Most students demonstrated at least L4 work, and only in a few cases less than L4.

Exams
• All three were implemented successfully with the go-kart the least problematic.
• In two of the cases there were difficulties in accessing the students’ work due to the range of file formats used and incompatibilities between file formats and software.
• For one case there were not enough computers for students to create their evidence portfolios.
• In all cases the assessment task was seen to be authentic (i.e. a relevant context for students, real tasks done in real time).
• In two cases task rubrics were developed for marking that resulted in highly correlated results for the two assessors but not with the teacher. In the other case only the course outcome scales of achievement were used and there was no significant correlation in marks.
• In one case this type of assessment was ‘part of the culture’ and was the best accepted by the students with the least difficulties and well correlated results. The other two cases required the teachers to provide a lot of instruction and varied in level of student efficiency and organisation.
• There is a fine balance between providing adequate structured scaffolding and providing sufficient flexibility to allow demonstration of achievement. The go-kart exam had very little scaffolding and thus marking required some ‘guesswork’ on student intention. The pool gate exam had a lot of scaffolding that enabled all students to end up with a representative product.

Reflective Portfolios
• There were no technical difficulties and the only logistical difficulty was a lack of time for the template-based slideshow
• All student work was saved on the school’s server and later transferred to an optical disc. In one case there were file size issues for the server.
• In all cases there was little, if any, correlation between the two assessors marks and none between them and the teacher.
• It was important that students were able to connect the value of the portfolio to later practical work. In one case this did not occur with a lot of written work that students could not see directly relating to practical outcomes.
• In all cases a lack of time was a serious constraint on the value of the assessment and the opportunity for students to represent their capability. However, in two cases the students were still positive about this approach to assessment and definitely preferred it to a written exam.
• Powerpoint portfolios are easiest to collect and to mark; all files in one folder, then uploaded into MAPS.
• Tasks-based folders, with work submitted in a number if formats were difficult to manage and mark, particularly as no student had successfully linked all the required items to a slideshow.
• Insisting on a particular structure given in a template is easiest to mark.

6.5.3 Summary of Overall Findings

This section provides a synthesis of the findings across both courses.

Manageability
• There were few, if any, technical or logistical difficulties for production or performance tasks exams in both courses.
• There were some logistical difficulties in some cases for both courses in managing longer period exams or portfolios (WA, AA, MA, GE, VE, WE& YE). This was usually concerning the management of student time to ensure they completed all requirements and in at least one case managing student access to computers.
• Portfolios were easier to manage with more structured templates, particularly web-based (e.g. WA & HE), with adequate time allocated.
• There were some difficulties in markers accessing student work where a large range of digital formats permitted (PA, XA, & YE). In one case (XA) all work was converted to PDF.
• In AIT a lack of student experience and ICT capability was an issue in a few schools (AA, WA & PA). In the case of PA this varied depending on student choice of software platform.

Technical
• Student work was collected either on a class DVD or USB drives and submitted either to a server or directly to the USB drives (one AIT and one Eng). There were no difficulties with these processes in any of the cases.
• In both courses some degree of maintenance was required to prepare student work for markers with the least for template-based web format output. The most was required for three AIT classes that used Flash or Powerpoint for reflective portfolios and for two Eng classes that used proprietary file formats for such as CAD drawings. For the AIT classes many students had linking errors. When the student work was in PDF, web or Flash format with errors removed, it was easily accessed by the markers.
• In Engineering a lack of computers and large file sizes for movies were issues in two cases.

Functional
• For all the AIT cases and for all but one of the Eng cases the students readily perceived the assessment task(s) to be authentic and meaningful. In all cases they preferred this to the alternative of a written exam. In all cases the teacher perceived the assessment to be more authentic than a paper-based exam.
• In one AIT case (CA) and one Eng case (WE) there was a substantial amount of written response work that students did not perceive to be relevant.
• In three AIT cases and four Eng cases the task was structured permitting a good range of levels of achievement to be demonstrated, in the other six cases this was not the case with either too little (XA, IA) or too much (AA, CA, WE, ME) structure.
• In many of the cases a lack of time limited the opportunity to demonstrate higher-level achievement.
• Where external assessors used a marking rubric developed from a standards-based framework almost always the result was highly correlated marks and rankings. For Eng the more the
rubric was constructed specifically for the task the more likely of highly correlated marks and rankings.

- In five of the AIT cases and all of the Eng cases there was little or no correlation between the teacher’s marks and ranking and the external assessors’ marks and rankings. This was even the case where in Eng the teacher used the same marking rubric as the external markers while in all AIT cases this did not occur. In most cases this appeared to be due to different interpretations of the outcomes and teachers incorporating background knowledge about the students and/or the implementation of the task. There were no discussions between the teacher and markers in any cases about how to assess. In two Eng cases in particular teachers clearly had access to more information about the student performance. This was particularly the case where the task was part of a larger sequence of tasks.

- In most AIT cases inadequate documentation and annotation by students limited achievement. Further this led to differences between external assessors due to interpretation of student intention. In most cases increased scaffolding and control of time use may have partly addressed this issue.

- The comparative pairs method of marking was successfully implemented for a short performance exam in AIT with resulting highly reliable scores but questions remain concerning the validity of results with variations for particular students.

**Pedagogic**

- Typically in both courses students liked doing practical work and disliked documenting work or answering question.

- In almost all cases for both courses the assessment matched general pedagogy for the course and was viewed positively by students.

- Assessment tasks worked best where the approach was familiar to the students with in one Eng case this being part of ‘the culture’. In reflective portfolios this tends to be the case due to the nature of the form of assessment whereas in the exams this depends on the design of the task(s). In all the AIT and all but one (WE), where there was too much written work, of the Eng reflective portfolios this was clearly the case. For both courses in all the production and performance task(s) exams this was also the case apart from one component of an AIT exam (CA) where there was a lot of written response work.

- In both courses most students demonstrated at least L4 with only a few in AIT judged at L5 and only a few in Eng below L4.

- In AIT the quality of work was low probably largely as a result of the capability of the students.

- In many AIT cases students liked learning ICT skills through the assessment. Particularly true in portfolios.

**6.6 Conclusions and Recommendations**

In this section overall conclusions from the study are stated with related recommendations discussed.

Each of the thirteen cases across the two courses involved different assessment tasks implemented in a unique fashion by the teacher in the school. An important finding of the study was that in all these cases there was sufficient technical knowledge and skill and IT infrastructure stability within the school to ensure every assessment task could be completed to an acceptable level. In at least two schools (AA & WA) a lack of student technical knowledge and skill limited the value of the assessment task(s) for those students but was implemented adequately by the teachers involved.

There was only one case where there was inadequate access to computer workstations and this was for an Engineering Studies class in a non-government school where most computer workstations were not close to the workshop within which the activity was taking place. This may be an issue for non-IT
courses where computers are often in a different location to the main performance (e.g. Engineering workshop).

In some cases there were particular problems associated with file formats. In most instances the problems concerned providing assessors access to the files although in one case the size of the files for the school server was a problem. It was found to be most efficient to either limit file formats to HTML, PDF, JPG, GIF or AVI or convert files to these formats. It is also likely to be necessary to limit the size of media files. Finally, where possible, it is easier for assessors if a student’s work is combined into a single file but this is dependent on the content types.

There appears to be a fine balance between providing adequate scaffolding and allowing students to demonstrate understanding by making choices. In this study there appeared to be a tendency towards too much structure, except for two AIT cases where there was judged to be an inadequate amount. However, greater structure did assist the marking process and therefore there may be a danger of unnecessarily stifling the opportunity for students to demonstrate understanding through over structuring tasks for marking convenience.

In most cases relatively detailed marking criteria were developed for the assessors based on the task(s) itself and the scales of achievement standards framework for the course. This was represented as a rubric. The study found that the results of assessor marking tended to be more highly correlated where the rubric included more information specific to the assessment task. Therefore it is recommended that to support highly reliable external marking assessment rubrics need to be specifically created for the assessment task but based on a standards framework. Further consensus discussions between markers are required to ensure consistency of meaning of criteria.

The study found that in nearly all cases there was little or no correlation between the results of marking by the teacher and the assessors. The lack of agreement in marking may be due to teachers taking into account background knowledge of the students or tasks. This may imply that teachers should not mark the work of their own students. At the least it is recommended that to support highly reliable marking by teachers considerable training is needed.

The study found that assessment tasks worked best where the approach is familiar to the students to the extent of being part of ‘the culture’. For the reflective portfolios this tended to be the case due to the nature of the form of assessment whereas for the exams this depended on the design of the task(s). However, for reflective portfolios this may not occur if there is too much written work that is perceived by the students to be irrelevant. In all the AIT classes and all but one Engineering Studies class’ reflective portfolios, students indicated that the approach was consistent with the course. For that one case the students judged that there was too much written work compared with that typical for the course. For both courses in all the production and performance task(s) exams the approach was familiar to students and found to be consistent with the pedagogy of the course.

### 6.6.1 Constraints

The study identified relatively few constraints to the use of the digital forms of assessment implemented in the sample of schools. The main constraint was clearly the amount of time students had to complete work. In both courses a lack of time was particularly a constraint for reflective portfolios. In one Engineering Studies case the short school period lengths reduced the engagement of students with the task.

Constraints concerning ICT were only evident in two cases. In one Engineering case the school’s file server was too small for the file sizes generated by students. In one Engineering case there were inadequate number of computer workstations.

There were some constraints in AIT connected to the capability of students. In AIT a lack of student ICT skills limited student engagement with tasks. In AIT a lack of organisation and literary skill limited student demonstration of understanding.

### 6.6.2 Benefits

Overall the study found that the benefits outweighed the constraints. In particular in both courses student responses were generally positive in all cases typically due to the practical nature of the work
and relevance to their interests. They certainly preferred this to paper-based exams. Most importantly in all cases the assessment task(s) could be completed at school with school equipment and be at least authenticated by the teacher.

There were slightly different benefits across the three general forms of digital assessment. Reflective portfolios generally permitted students to address a greater range of outcomes and demonstrate a greater range of knowledge and skills although in a few cases a lack of time limited this opportunity. Production exams generally permitted students to address a limited range of outcomes and demonstrate a reasonable depth of knowledge and skills dependent on the task and time. Performance task(s) exams generally permitted students to address one or two outcomes and demonstrate a limited depth of knowledge but a reasonable level of skill dependent on the task and time.
References


Appendices

Appendix A: Design Brief for AIT

A design brief was developed for the Applied Information Technology course based on the initial literature review.

Summary of Assessment Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Extended Production Exam</strong></td>
<td>• Full use of technology process (longer the better),</td>
<td>• Difficult to assess the higher levels</td>
</tr>
<tr>
<td>(Focuses on one common assessment task and incorporates a full technology process. Focus on technology process, skills and enterprise. Centres on one major scenario rather than a series of smaller scenarios.)</td>
<td>• Open-ended to allow students to show what they can do</td>
<td>• Lack of time limits what students demonstrate</td>
</tr>
<tr>
<td></td>
<td>• Range of skills (e.g. management, technical)</td>
<td>• Time consuming to mark (software access)</td>
</tr>
<tr>
<td></td>
<td>• Quality of design as evidenced in output</td>
<td>• Expensive to organise</td>
</tr>
<tr>
<td><strong>Focused Performance Tasks Exam</strong></td>
<td>• Short</td>
<td>• No overview of Technology Process,</td>
</tr>
<tr>
<td>(A range of small focussed practical tasks that students submit after working under exam conditions at a computer. Each task is connected with a small scenario.)</td>
<td>• Relatively cost-efficient</td>
<td>• Needs many items and choices to cover range of student experience,</td>
</tr>
<tr>
<td></td>
<td>• Focus on specific outcomes</td>
<td>• Time consuming to mark (software access),</td>
</tr>
<tr>
<td><strong>Digital Portfolio</strong></td>
<td>• Work already done</td>
<td>• Limited time makes it more difficult for students to demonstrate the higher levels – indicative achievement only,</td>
</tr>
<tr>
<td>(A portfolio of work by a student is assessed by two examiners. Specify parameters for submission – form, structure, range of samples required.)</td>
<td>• Easy to cover all outcomes</td>
<td>• Software requirements for schools.</td>
</tr>
<tr>
<td></td>
<td>• Relies on output and documentation</td>
<td>• Large quantity takes a long time to mark</td>
</tr>
<tr>
<td></td>
<td>• Difficult to authenticate</td>
<td>• Difficult to authenticate</td>
</tr>
</tbody>
</table>

Extended Production Exam

This is a common assessment task, done under exam conditions at a computer, and incorporating a full technology process (similar to MSE example). The focus of activity is on the technology process, skills and enterprise. The assessment task would centre on one major scenario that would enable students to select the most appropriate software and medium of presenting information. The task would need to have severe constraints to enable students to complete it in a reasonable time.

There would be a particular focus on outcome 3 and to a lesser extent outcome 1. This has limitations in terms of what can be produced and to what level in a contained/limited time.

In order to be suitable for the range of units it could not be software specific, however students could be given question options. The task is built around the same parameters. Could use a common design situation. Need to provide generic formats and concepts, how to submit files (attachment, ftp, burn on CDROM etc.). Purposely left in note form?

Addressing of Outcomes

All outcomes are addressed but only Outcomes 1, 3 and 4 would be addressed adequately.
Outcome 1: Full technology process implemented but in limited form.
Outcome 2: Limited by the content surrounding the scenario.
Outcome 3: Limited opportunity to demonstrate high level creativity but enough to cover all aspects.
Outcome 4: Reasonable opportunity to demonstrate both understanding and application.

Nature of Items
Students would demonstrate knowledge and skills through creating design documents, prototype solutions in electronic form, and responding to a series of both closed and open questions. All responses will be in electronic form with alternatives to continuous prose utilised through scaffolded tables, dot points, diagrams, etc.

Nature of Stimulus Materials
Suitable stimulus materials would be in electronic format and describe the scenario providing adequate information for students to search to complete all aspects of the assessment task. This could include articles, audiovisuals, diagrams, images, design or ICT screen shot samples or exemplars. In essence the student is given a full research report on the scenario.

Constraints/Resource Implications
Items and materials must suit students who may have studied different combinations of units and a variety of contexts. Copyright permissions will be required for stimulus materials subject to copyright restrictions or where, for example, ICT proprietary software screen shots are used. It is not anticipated that colour printing of stimulus materials is required. The nature of these outcomes and items will suit a computerised exam format that includes static and interactive elements.
PROS: Full use of technology process (longer the better), open-ended to allow students to show what they can do, range of skills (e.g. management, technical), quality of design as evidenced in output,
CONS: Difficult to assess the higher levels, lack of time limits what students demonstrate, time consuming to mark (software access), expensive to organise.

Sample Instructions
Investigate, devise, produce and evaluate a computer-related technology (i.e. product, service, process, system, or environment) associated with the given scenario.

Description of Tasks
1) PROPOSAL: A proposal for the development of a technology including a definition of the industry and the problem.
2) DESIGN DOCUMENT: A design document for the technology that may include drawings, flowcharts and schematics.
3) PROCESSES PLAN: A plan of project activities, sequencing and logistics.
4) PROTOTYPE: A prototype of the technology indicating all significant components.
5) REFLECTION: Responses to questions that reflect on the processes and outcomes.
The output from all tasks will be collated into an electronic folder on a DVD and must be accessed through a browser such as Internet Explorer and may also call on readers of .doc, .ppt and .xls files or plugins for .pdf and .wav (etc.). Any other plugins or readers must be provided by the student or school.
Parameters of Tasks

(1) PROPOSAL
The student will develop a proposal to design and develop a technology (i.e. product, service, process, system, or environment) based on an aspect of the use of information technology in the industry that was investigated in the research report. The proposal should include:

- An outline of the human need or opportunity that is being addressed.
- The main objectives of the project.
- Briefly describe existing solutions that were identified in the research report.
- Explain what aspects of existing solutions need improving and thereby the criteria that could be used to evaluate the success of the project.
- Identify what resources may be required by the project, in particular hardware, software, and technical ICT skill.

The teacher will negotiate with the student(s) the scope within the budget and resources of the school/provider. Therefore the teacher should provide the student with detailed feedback on the proposal. The student may then use this feedback to modify aspects of the project that should then be reflected in the remaining tasks.

Students may work individually or in groups. If students work in groups then each team member is to participate in aspects of organisation, planning and production of the project. The scope of the presentation should be such that all group members can be fully involved. The teacher will oversee the allocation of students to groups to ensure equity.

(2) DESIGN DOCUMENT
The student will develop a design document for the technology that may include descriptions, photographs, drawings, flowcharts and schematics.

- Examples of early attempts which were subsequently improved should be included with explanation of the improvements.
- The document should include adequate information that would allow another skilled person to complete the project.

(3) PROCESSES PLAN:
The student will develop a brief plan of project management, activities, sequencing and logistics. This will include the need to develop the skills and understanding in applying appropriate pieces of hardware and software.

- Include all documentation developed to manage the project.
- Briefly document all resources applied to the project: hardware, software, materials and personnel.
- Include plans for the development of skills and understanding in applying appropriate pieces of hardware and software.

(4) PROTOTYPE:
The student will produce and provide a prototype or completed copy of the technology. This will be provided in one or more electronic files on the DVD.

- Students will apply at least two pieces of computer software in the project.
- Include technical and user manuals that are accessed through a web browser.

(5) REFLECTION:
The student will respond to a series of reflective questions on the processes and output of their activity. The content would include questions concerning:

- the technology process employed.
- the key-points of the design.
- design ideas, tools and techniques.
- justification of ideas and resources.
- comparisons with processes and techniques used in the industry.
**Authentication**

The teacher will monitor the progress of the production. The teacher’s assessment of the project may be included in the scheme of school-based assessment.

The submission of the project DVD should be accompanied by a certificate signed by the school principal, the supervising teacher, and the student verifying that all work submitted is the student’s own work except where clearly referenced and completed under the supervision and guidance of the teacher.

A brief written record of each student’s progress throughout the project must be kept by the teacher. This may be requested in exceptional circumstances where the examiners require further information. This record should be retained in the school together with assessment records.

**Focussed Performance Tasks Exam**

A range of small focussed practical tasks that students submit after working under exam conditions at a computer. Each task is connected with a small scenario. A scenario would need to be used that enables students to select the most appropriate software and medium of presenting the information. The task would need to have certain constraints built into it to enable students to complete it in a reasonable time.

There would be a particular focus on outcome 3 and to a lesser extent outcome 1. This has limitations in terms of what can be produced and to what level in a contained/limited time.

In order to be suitable for the range of units it could not be software specific, however students could be given question options. The task is built around the same parameters. Could use a common design situation. Need to provide generic formats and concepts, how to submit files (attachment, ftp, burn on CDROM etc.). Note format OK?

**Addressing of Outcomes**

*Suited to Outcome 1 (Aspects 2 and 3) and Outcome 3.*

Outcome 1: Only aspects of the technology process implemented in limited form.
Outcome 2: Limited by the content surrounding the tasks.
Outcome 3: Limited opportunity to demonstrate high level creativity but enough to cover all aspects.
Outcome 4: Not suited but could ask questions in response to some tasks.

**Nature of Items**

Students would demonstrate knowledge and skills through completing the tasks and responding to questions about the tasks. The performance item(s) should enable all candidates to apply a broad range of technical skills (increasing in complexity) independent of a range of contexts. Candidates could complete a range of small focussed practical tasks on a computer or produce and submit at least a prototype technology based on a design proposal. Each item/series of items is connected with a small scenario. The item(s) will specify the nature (form, structure, range of target audiences required), the type and number of digital products or prototype files (within specified constraints) to be included and the file formats required (dependent on further research). Some tasks require the performance of technical and organisational skills using hardware and software. Question scaffolding will be used to encourage responses at the higher levels. Instructions to candidates will specify how to submit files (i.e. burn onto CDROM).

**Nature of Stimulus Materials**

This could include articles, audiovisuals, diagrams, images, design or ICT screen shot samples or exemplars. In order to be inclusive of the range of contexts and software candidates may have studied, items cannot be software specific. However, optional items could be used. This may entail using generic formats and concepts.
Constraints/Resource Implications

Items and materials must suit students who may have studied different combinations of units and a variety of contexts. Copyright permissions will be required for stimulus materials subject to copyright restrictions or where, for example, ICT proprietary software screen shots are used. The nature of these outcomes and items will suit a computerised exam format that includes static and interactive elements. Each candidate will need access to a computer. Parameters will be needed to standardise the exam administration, such as the hardware and software required.

A variety of performances across contexts may be difficult to compare. The scenario would need to allow candidates to select the most appropriate software and medium for presenting the information. Some constraints would need to be built into the scenario so it can be done in a reasonable time and assess higher levels in a contained/limited time. Output file formats would need to be standardised which may limit the variety of performance. This is dependant on the selected computerised delivery format, specified parameters, and resource availability.

PROS: Short, cost-efficient, focus on specific outcomes,
CONS: No overview of Technology Process, needs many items and choices to cover range of student experience, time consuming to mark (software access), limited time makes it more difficult for students to demonstrate the higher levels on the scale – looking at indicative achievement only, software requirements for schools.

NOTE: Simulated performance tasks (online) could be used with a range of small focussed simulated tasks that students submit online after working under exam conditions at a computer through a browser. Each task is connected with a small scenario.

Digital Portfolio

A portfolio of work would be submitted by the student to be assessed by a panel of assessors. This would specify certain parameters for submission of the portfolio – form, structure, range of samples required. Electronic files could include a video presentation by the student and other types of audiovisuals. Need to clearly identify files and software and locations.

Addressing of Outcomes

All outcomes are addressed adequately.

Nature of Items

A digital portfolio is submitted separately on a CD/DVD. The digital portfolio will allow candidates to apply a broad range of skills (increasing in complexity) independent of the range of contexts studied within the school. The digital portfolio specifications will specify the nature (form, structure, range of target audiences required), the type and number of examples to be included and the browser/file formats required and will be authenticated by teachers and principals.

Constraints/Resource Implications

The digital portfolio specifications must suit students who may have studied different unit combinations and a variety of contexts. While schools will require access to computer hardware and software, digital video cameras, and editing facilities, etc, these resources and facilities are required to effectively deliver this course of study. Candidates or schools will provide the CD/DVD.

PROS: Work already done, easy to cover all outcomes,
CONS: Relies on output and documentation, large quantity take a long time to mark, difficult to authenticate.
Sample Instructions

Students will complete and submit a ‘Body of Evidence’ Portfolio. They may then sit an examination that will include items that require them to reflect upon their own portfolio and respond to other stimulus material.

Students are required to select digital artefacts and process documents that demonstrate their highest achievement in the creative and conceptual realisation of the application of information and communications technologies (ICT) to address human needs. The emphasis is on quality not quantity. It contributes to the assessment of the student’s ability to use and evaluate the application of ICT relevant to personal, community, business and industry settings. The portfolio provides evidence for experience in which they have developed an understanding of, and practical skills in the use of ICT relevant to these settings.

Each digital file should include an annotation indicating how it was created, the format of the file, and an explanation of the purpose and uses of the file.

Description of ITEMS

The digital portfolio should include items organised according to the following structure with at least TWO items in each sub-area. The two items should be associated with different activities/projects and illustrate different aspects of the sub-area.

TECHNOLOGY PROCESS: Annotated process documents that illustrate how in creating or modifying products using ICT the student,
(a) investigates ideas and evaluates alternative designs and proposals;
(b) devises, communicates and evaluates proposals and design plans in appropriate forms; or
(c) implements and evaluates production processes and strategies to manage resources efficiently.

UNDERSTANDING ICT: Annotated digital artefacts and/or process documents that illustrate the student’s,
(d) understanding of the ICT-related concepts, standards and terminology required to select and use appropriate computer software and hardware to achieve information solutions;
(e) understanding of the standards and conventions when selecting and using ICT systems to achieve information solutions; or
(f) understanding of the management, processes, functions, types and relationships between the components of ICT systems required to achieve information solutions.

QUALITY OF INFORMATION SOLUTIONS: Annotated digital artefacts and/or process documents that illustrate the student’s,
(g) application of accepted standards and conventions to create or modify various forms of information solutions;
(h) application of efficient skills, techniques and processes in the use of ICT systems to develop information solutions; and
(i) application of enterprising capabilities while exploring alternatives and working to achieve information solutions.

ICT IN SOCIETY: Annotated digital artefacts and/or process documents that illustrate the student’s,
(j) understanding of the cultural beliefs, values, abilities and ethical position that can impact on the use of ICT;
(k) understanding of the benefits and consequences of ICT use in different contexts and how this relates to beliefs, values, abilities and ethical positions; and
(l) understanding of the consequences of technological developments on social, personal, cultural, physical, economic and ethical structures and environments.
Parameters of PORTFOLIO

The teacher will advise and set parameters within the budget and resources of the school/provider. The teacher will provide examples of the standard expected for portfolio items and discuss with students the variety of work that could possibly be included. Typically these items will be generated from learning activities organised by the teacher at school including some which may be drawn from school assessed tasks, while others may be organised by the student. An item should not be a complete school assessed task. Students may draw on work completed individually or in groups but may only include work they have contributed to and must indicate their personal contribution.

Students shall be given appropriate access to hardware and software for at least 30 hours for the modification and annotation of digital items for the portfolio. It is envisaged that over a two-unit period students would spend 30-40 hours collating their portfolio in addition to the work they are required to do in the units and therefore it is anticipated that much of this will occur outside normal class times (e.g. annotations may be written at home).

The entire digital portfolio will be collated onto a CD/DVD-R and must be accessed through a standard browser such as Internet Explorer and may also call on readers of .doc, .ppt and .xls files or plugins for .pdf, .swf, .mov and .wav. Any other plugins or readers must be provided on the CD/DVD-R.

The portfolio must, as a minimum, contain examples of the student’s own work to cover the FOUR areas (outcomes), each comprising THREE sub-areas (aspects) listed below. Students must not include more than TWO items for any aspect of an outcome. A portfolio template comprising HTML files will be provided within which to collate the portfolio.

Following are some examples of the types of items that may be included for each aspect.

1) TECHNOLOGY PROCESS

The portfolio should include process documents that illustrate how in creating or modifying products using ICT the student employs an appropriate technology process. In each area the focus is on examples of the student’s work that illustrate their highest achievement. Examples of items that may be included are as follows.

(a) The student investigates ideas and evaluates alternative designs and proposals.
   • A document that shows two alternative designs and is annotated to indicate how/why the designs were created, in terms of the human need or opportunity being addressed, and why one was selected.
   • A document indicating the search strategies used to find information from multiple sources and how this information was selected, analysed and subsequently used.
   • A document illustrating the use of a strategy to investigate ideas for a project along with annotations evaluating this strategy.

(b) The student devises, communicates and evaluates proposals and design plans in appropriate forms.
   • A document outlining a proposal that includes information concerning the human need or opportunity being addressed, the objectives etc.
   • An electronic presentation of a proposal including early attempts which were subsequently improved with an explanation of the improvements.
   • A digitally created design plan for the creation or modification of a technology that includes an explanation of the key features of the design and why/how these were developed. The design plan may include descriptions, photographs, drawings, flowcharts and schematics.
   • A document outlining an evaluation of a proposal or design plan including the criteria used in the evaluation, the results of the evaluation, and what aspects need improving and how this would be achieved.
(c) The student implements and evaluates production processes and strategies to manage resources efficiently.

- A document identifying the resources (e.g. hardware, software, technical ICT skill) required to implement production processes and strategies for the creation of a school newspaper including a plan for its efficient management.
- An electronic presentation of the implementation of production processes and strategies for the creation of a digital recording of the school band.
- A document outlining the planning and implementation of a survey of students on the design of a new school uniform including the management and analysis of the data.

(2) UNDERSTANDING ICT

The portfolio should include annotated digital artefacts and/or process documents that illustrate the student’s understanding of the application of computer systems to the creation and modification of information solutions (i.e. product, service, process, system, or environment). In each area the focus is on examples of the student’s work that illustrate their highest achievement. Examples of items that may be included are as follows.

(d) The student understands the ICT-related concepts, standards and terminology required to select and use appropriate computer software and hardware to achieve information solutions.

- An electronic slideshow of a Talking Book, saved as a movie, along with annotations explaining technically why particular design features were incorporated and why the particular software and hardware were selected.
- Three sound files created/modified using some audio editing software along with annotations explaining technically some of the operations that were used and the impact on file size, quality, portability etc.
- A document with three images of the same photograph taken by the student along with annotations describing the differences between the images as these relate to the ICT-related concepts, standards and terminology.
- A document with two images created by the student using digital tools along with annotations describing differences between the images in terms of colour, emphasis, proximity, symmetry and contrast.

(e) The student understands the standards and conventions when selecting and using ICT systems to achieve information solutions.

- A spreadsheet to present the accounts for the school’s student council with annotations explaining the use of standards and conventions.
- A document with two images created by the student using digital tools along with annotations judging differences between the images in terms of standards and conventions in colour, emphasis, proximity, symmetry and contrast.
- A document with three images of the same photograph taken by the student along with annotations describing how the images would be judged in terms of accepted standards and conventions when using digital image editing systems.

(f) The student understands the management, processes, functions, types and relationships between the components of ICT systems required to achieve information solutions.

- A document outlining a plan of project management, activities, sequencing and logistics showing skills and understanding in applying appropriate pieces of hardware and software.
- A digital video of a prototype simulated robotic system to improve the control of processes (lighting, heating, food) for the fish tank in the library of the school. This includes annotations of how the components of the system combine to achieve the solution.
• A document evaluates the use of a computer system in a local video hire store in terms of the management, processes, functions, types and relationships between the components of the ICT system.

(3) QUALITY OF INFORMATION SOLUTIONS
The portfolio should include annotated digital artefacts and/or process documents that illustrate the student’s capabilities in applying knowledge, skills, and enterprising attitudes to the creation or modification of information solutions (i.e. product, service, process, system, or environment). In each area the focus is on examples of the student’s work that illustrate their highest achievement. Examples of items that may be included are as follows.

(g) The student applies accepted standards and conventions to create or modify various forms of information solutions.
• A small business yearly budget created as a spreadsheet, along with annotations explaining how accepted standards and conventions were applied.
• An electronic slideshow of a Talking Book, saved as a movie, along with annotations explaining how accepted standards and conventions were applied.
• A music sound file created using MIDI technology suitable for delivery through the iTunes music store.
• A digital file for a computer controlled sewing machine that produces a design for a skirt.

(h) The student applies efficient skills, techniques and processes in the use of ICT systems to develop information solutions.
• A flat-file database created to illustrate the main data required to enact an e-commerce transaction, along with annotations explaining one or more set of processes that would be used to ensure the efficient use of this information solution.
• An electronic slideshow of a Talking Book, saved as a movie, along with annotations explaining one or more particular skills, techniques and processes that were used to improve the efficiency and value of the information solution.
• A claymation created as a Flash animation and presented as a Shockwave file along with annotations explaining one or more particular skills, techniques and processes that were used to improve the efficiency and value of the information solution.
• A flash animation using graphics and sound with annotations explaining one or more particular skills, techniques and processes that were used to improve the efficiency and value of the information solution.
• A document with three images of the same photograph taken by the student along with annotations describing how the photograph was edited to create each of the images and what the purpose of each was.
• A prototype information education system for a museum based on RFID technology.

(i) The student applies enterprising capabilities while exploring alternatives and working to achieve information solutions.
• An online school newsletter that automates the layout and submission of articles, along with annotations explaining the advantages over alternative solutions.
• An electronic slideshow of a Talking Book, saved as a movie, along with annotations explaining how this represents the application of enterprising capabilities.
• A database system, to be used by a lawn-mowing home business, that efficiently organises bookings.
• A document outlining a plan to improve the canteen lunch ordering system using an online booking system, credit card payment and automated delivery system.
(4) ICT IN SOCIETY
The portfolio should include annotated digital artefacts and/or process documents that illustrate the student’s understanding of the impact of information solutions on society and the influence of aspects of society on the use of ICT. In each area the focus is on examples of the student’s work that illustrate their highest achievement. **Examples of items** that may be included are as follows.

(j) The student understands the cultural beliefs, values, abilities and ethical position that can impact on the use of ICT.
- A process document that shows that the student considered beliefs, values, abilities and ethical positions in the design of a database for use in a medical practice.
- An digitally created poster that explains how to use an online voting system with annotations indicating how an understanding of cultural beliefs, values, abilities and ethical position were used in the design.
- A document setting out the design of a user interface for an online ordering system for lunches from the school canteen indicating how an understanding of cultural beliefs, values, abilities and ethical position were used in the design.

(k) The student understands the benefits and consequences of ICT use in different contexts and how this relates to beliefs, values, abilities and ethical positions.
- An electronic slideshow of a Talking Book, saved as a movie, along with annotations explaining how the beliefs, values, abilities and ethical positions of stakeholders such as parents, community groups and students were considered in the design and development.
- A document presenting the results of a survey, conducted by the student, of the impact of the ownership of 3G mobile phones by students and how this relates to the beliefs, values, abilities and ethical positions of students, parents and teachers.
- A document analysing the benefits and consequences of a computer controlled milking system and how this relates to beliefs, values, abilities and ethical positions.

(l) The student understands the consequences of technological developments on social, personal, cultural, physical, economic and ethical structures and environments.
- A process document that outlines the evaluation of the use of a computer controlled sewing machine to produce new swimwear in terms of the impact on social, personal, cultural and economic environments.
- A process document that explains how research into the impact of RFID technology was used to inform the design of a prototype museum education information system.
- An online marketing tool with annotations indicating the way in which an understanding of the social, personal, cultural, physical, economic and ethical structures and environments has informed its design.

**Submission and Authentication**
The portfolio is to be submitted on a single disc (DVD or CD).

School candidates must submit their portfolio through the school and must complete a declaration form to declare that:
- The work submitted is the student’s own.
- The work submitted was completed as a part of the student’s studies.
- The work is not dangerous or offensive.

Any portfolio submitted without a completed declaration form will not be marked. Declaration forms will be sent to schools and candidates. The candidates will be provided with a student number to be referred to in each portfolio item (e.g. header/footer). It is the candidate’s responsibility to ensure that each item submitted is labelled securely with their student number. A candidate’s name, or the actual names of persons associated with the candidate’s school or family, or the school name, must not appear on any item.
The submission of the portfolio CD/DVD should be accompanied by a certificate signed by the school principal, the supervising teacher and the student verifying that all work submitted is the student’s own work except where clearly referenced and completed under the supervision and guidance of the teacher. A proforma will be provided for this verification procedure. The teacher will monitor the progress of the portfolio during its creation.

Students will include references to all information sources they have used. Students will identify the nature and source of any assistance they have received and information they have used. Where components of a portfolio item were undertaken by some other person or agency the contribution must be documented in the annotations for that item. Students will not be given credit for actual work completed by others. Justification for, and of, such work will be recognised in the marking process.

An ongoing backup of the portfolio must be kept by both the student (off-site) and the school (on-site) and be provided if required. Ultimately it is considered the responsibility of the student to maintain a backup copy but a final copy should be retained in the school together with assessment records. In addition, under exceptional circumstances where the examiners require further information, copies of learning programmes and learning activity instructions may be requested from the teacher.
Appendix B: Design Brief for Engineering Studies

The three digital formats examined in this research are: Extended production Examination; Focussed Performance Task; and Electronic Portfolio.

Extended Production Exam

This is a task done under exam conditions including both real and virtual aspects, which together cover the spectrum of production activities. So students will respond directly to questions on the computer, and other responses will involve manipulative activities in a workshop. These latter activities will be digitally recorded and so form part of the evidence of student performance.

The assessment task will centre on the students reaction to a design context, scenario or problem. The student activity will be guided and limited to ensure that it is completed within the set timeframe.

All outcomes (Engineering process, Engineering understandings, Engineering skills, and Engineering in Society) can be assessed in this type of examination but given the limited timeframe, not all aspects will be addressed. The question guidance will focus on those elements that require assessing.

Nature of Items

Items may include design documents, prototypes, and questions may include graphics, open and closed, short and extended response. Directions will be also provided for the specific workshop activities students are to do in relation to the task. These activities will be digitally recorded to provide part of the evidence.

Nature of Stimulus Materials

Suitable stimulus materials will be in electronic format and describe the scenario, providing adequate information for students to complete all aspects of the assessment task. This could include a video, diagrams, text, images or animations.

Sample Instructions

If a school was specialising in the context of structures and mechanical systems, a problem could be presented to students to design a shelf to span a specified distance and support a set weight. The sequence of questions could follow the design process: drawing out student’s knowledge of a range of materials, creative and analytical exploration of a range of design options, application of the principles of structures and forces, and the design of tests to ascertain factors of safety. The practical element could include the production of a joining system or a test apparatus.

In this way the student is proceeding through a production exercise in a guided manner, targeting specific knowledge and skills, and including a practical element, which is so essential to the Engineering course.

Focussed Performance Tasks

Focussed performance tasks are short tasks that students submit after working under examination conditions in a workshop and/or at a computer to digitally record workshop output. The tasks develop from an engineering scenario, and the questions or instructions focus the student’s work on a specific task related to the scenario. The task would have certain inbuilt time constraints to enable the student to complete it in a reasonable time.

Student performance on the tasks may be through simulation on a computer or through a real performance task, which is then recorded digitally in order to provide electronic evidence of performance. The task may provide for a combination of virtual and real activities. This assessment task would focus on:

1. Outcome 1: Engineering Process,
2. Outcome 2: Engineering Understandings and
Each Performance Task would focus on just one or two aspects of the outcomes.

**Nature of Items**
Students would demonstrate knowledge and skills through completing the tasks and responding to questions about the tasks. The performance item(s) should enable all candidates to apply a broad range of technical knowledge and skills (increasing in complexity) regardless of the contexts.

Candidates could complete a range of small focussed practical tasks on a computer, or produce and submit evidence of performance.

**Nature of Stimulus Materials**
Suitable stimulus material could include articles, audiovisuals, diagrams, images, engineered products or exemplars.

**Sample Instructions**
Items and materials must suit students who may have studied different combinations of units and a variety of contexts.

If a school wished to assess Outcome 4, *Engineering in Society*, by testing *Socially conscious engineering* content, students could be presented with a scenario which contrasted engineering production methods in a developing country with an abundance of cheap labour, and a more developed country. The performance tasks would test the student’s ability to apply engineering principles to contrasting contexts.

Alternatively if the focus was on Outcome 1, *Engineering process*, a scenario could relate to the use of solar power in remote farming communities. The performance tasks would test the student’s skills in investigating and devising.

**Electronic Portfolio**
A portfolio of work would be submitted by the student to be assessed by a panel of assessors. The brief for the portfolio would specify certain parameters for submission of the portfolio – form, structure, and range of activities for which evidence is required. Electronic files including text, video, digital still pictures, CAD, and scanned sketches and images, could be included.

All outcomes are addressed adequately.

**Nature of Items**
A digital portfolio is submitted on a CD/DVD. The digital portfolio will allow candidates to provide evidence of a broad range of skills independent of the contexts studied within the school.

**Nature of Stimulus Materials**
No stimulus material is required for this form of digital assessment.

**Sample Instructions**
Students will complete and submit a Portfolio. They may also complete an examination that requires them to reflect upon their own portfolio and respond to other stimulus material.

Students are required to select exemplars of their work that demonstrate their highest achievement in specified outcome aspects. The folio specifications may not include aspects that are adequately assessed elsewhere throughout the course. The emphasis is on quality not quantity.
Appendix C: Standards-based marking tool
Appendix D: Comparative pairs marking tool