2011

An Investigation into the feasibility of using digital representations of students' work for authentic and reliable performance assessment in applied information technology

David G. Miller
Edith Cowan University

An Investigation into the Feasibility of using Digital Representations of Students' Work for Authentic and Reliable Performance Assessment in Applied Information Technology

by

David George Miller

BSc (Hons), PGCE, MEd, Grad Dip IT

This thesis is presented in fulfilment of the requirements for the

Degree of Doctor of Philosophy

Faculty of Education

Edith Cowan University

Perth

Western Australia

June 2011

Supervised by

Associate Professor C. P. Newhouse

USE OF THESIS

DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

(i) incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education.

(ii) contain any material previously published or written by another person except where due reference is made in the text; or

(iii) contain any defamatory material.

I also grant permission for the Library at Edith Cowan University to make duplicate copies of my thesis as required.

Signature: …………………………………….

Date: January, 2012

ACKNOWLEDGEMENT

I wish to take this opportunity to thank my principal supervisor, Associate Professor C. P. Newhouse, for his tireless support and advice over the last four years. Associate Professor Newhouse's willingness to read and re-read my work, his clarity and precision of thought and meticulous attention to detail have all been invaluable, in fact essential, to the completion of this thesis. I would also like to acknowledge Dr Jeremy Pagram for his advice, comments and proofreading of the final manuscript.

ABSTRACT

The assessment of student performance in areas such as drama, physical education, art and Information Technology (IT), does not lend itself to traditional, paper-based testing methods. In these domains, much emphasis is placed on the acquisition and demonstration of practical skills and these may be difficult, if not impossible, to measure by scores on theoretical, written assessments. Alternative forms of assessment, which are both valid and reliable, need to be devised for the practical aspects of these subject areas. The capture, in digital form, of students' work, may allow the development of authentic forms of summative, high-stakes assessment with high reliability. This study investigated the digital capture of aspects of the practical performance of students in the senior secondary course of Applied Information technology (AIT), across seven high schools in Western Australia. Two forms of assessment were investigated; a reflective process digital portfolio and a computer based production examination.

This study formed part of a larger project investigating the feasibility of using digital representations of students' performances for authentic and reliable assessment in senior secondary school courses. This study only focussed on the AIT course, one of the four courses investigated, and only the first 'proof of concept' phase of the three developed by the main project. An ethnographic, action research methodology was employed, using qualitative and quantitative data collected and compiled into multiple case studies. The main sample comprised 115 students in eight classes across seven schools, resulting in seven case studies. These students completed a digital portfolio over a four-week period and a computer based practical/production examination over three hours. The examination also included a response questions section. Portfolios were scored by summation of partial marks according to a marking rubric; examinations were scored similarly and, in addition, for a subset of students, by a method of multiple comparisons of pairs. For each method of marking Rasch modelling analysis was conducted to investigate the reliability of scoring.

The researcher observed all classes a number of times as the two forms of assessment were completed. Subsequently, students completed a 70 question survey designed to elicit their viewpoint on each of the two forms of assessment, their attitude towards, ownership of and frequency of use of digital technologies, and a self-assessment of their own computer skills. For each class, following the examination, one or two small groups of students took part in a short forum where they responded to a series of questions about the examination. Additional data were obtained from the teachers, who took part in interviews and submitted their own set of students' results. Interviews with the external assessors of the students' work completed the range of data sources. Data were analysed for each case study and for the combined sample using both qualitative and quantitative techniques. The results of data analysis were interpreted through a feasibility framework developed from one used by Kimbell et al. (2005) in the e-scape project. This supported an investigation of the manageability, technical feasibility, functional operation (validity and reliability), and pedagogical alignment of each form of assessment.

The digital portfolio was implemented in each class by the teacher while the examination was invigilated by the researcher and the teacher and facilitated by providing students with a USB flash drive containing all the resources required. For both forms of assessment, the task was developed by a team in the light of a situation analysis and was implemented in the students' normal laboratory using the workstations and software normally available. The study found that in general students viewed both forms of assessment very positively and were almost unanimous in their preference for computer-based assessment over written examinations. In general, teachers regarded the form and scope of the assessments favourably, though there were some criticisms of the examination theory response questions and marking key. Markers found the on-line system quick and easy to use for both the analytical and the comparative-pairs marking. Further the time taken was similar for both methods.

Although digital capture of students' performance was not without problems, the study concluded that the benefits far outweighed the constraints. Digital capture allowed authentic practises to be demonstrated, stored, transported and assessed analytically with high reliability for both the portfolio and the examination. Assessment by the method of multiple comparisons of pairs was shown to be a feasible alternative to analytical marking with good overall reliability even though for some work samples there was a wide disparity between the rankings generated by the two methods. It was concluded that a major explanation for this discrepancy was where there appeared to be a substantial difference between a student's technical skills and creative design skills.

Each form of assessment was found to have had relative strengths and weaknesses. The computer-based production examination, with its concise and structured format was implemented more consistently than the digital portfolio, where teachers' interpretations of the requirements differed widely. Work produced during the examination was entirely that of the student whereas for the portfolio, collaboration and assistance could not be discounted. The portfolio, with its broader scope and opportunity to demonstrate a wider variety of skills, was a more valid form of assessment than the examination. In the latter, the short time frame restricted tasks to a relatively low level in order that students from all classes could attempt and complete them. Further, the results of marking suggested that the portfolio was scored with greater reliability than the examination. There were some minor technical difficulties during the examination but none of these prevented completion. For the portfolio, the extended time frame meant that any technical difficulties could be resolved without impacting on the assessment process.

The study recommended that the current paper-based final examination be replaced by one of the two alternatives; a computer based production examination or a digital portfolio. The possible form and implementation of each method is set out in the light of the findings of the study. The study clearly demonstrated that either form of assessment could be readily implemented in schools and would be preferable to the current paper-based form that is not well aligned with the requirements of the course syllabus nor to the needs of students, teachers and the workplace.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEFINITION OF TERMS AND ABBREVIATIONS

- *Digital representations of student performances:* electronic files of students' work recorded as film, photographs, audio, text and/or graphics.
- *Extended Production Examination:* a task completed under examination conditions, incorporating a full range of processes. For example, the design, creation and appraisal of a product.
- *Focussed performance task:* a practical task completed under examination conditions and submitted in digital format.
- *Reflective process portfolio:* a collection, in digital form and according to a predetermined structure and sequence, of the work output *during* the completion of a task. Files might include; initial ideas, design sketches, reflective commentary, video and photographs.
- *Recorded interview:* a video or audio recording of the student's responses to a series of scripted questions and prompts designed to illicit the thinking processes connected with completion of a task.
- *Manageability of digital form of assessment:* pertaining to the practicalities of administration, collection and assessment of artefacts of student work in digital forms.
- *Technical facility of digital form of assessment:* concerning the extent to which existing technologies are suitable for adaptation to the purposes of assessment.
- *Pedagogy of digital form of assessment:* pertaining to the extent to which digital forms of assessment can support and enhance teaching and learning
- *Functionality of digital form of assessment:* concerning the validity and reliability of digital forms of assessment and their comparability with other methods of assessment.
- *MS*: Microsoft
- *USB*: Universal Serial Bus
- *MB*: Mega Byte
- *GB*: Giga Byte
- *PDF*: Portable Document Format

- *HTML*: Hypertext Mark-up Language
- *DVD*: Digital Video(Versatile) Disk
- *LAN*: Local Area Network
- *PHP*: General purpose scripting language for dynamic webpages
- *SQL*: Structured Query Language
- *SPSS*: Statistical Package for the Social Sciences
- *SD:* Standard Deviation

CHAPTER ONE

INTRODUCTION

Background to the Study

At the beginning of the current millennium, Prensky (2001) claimed that today's secondary students are digital natives in a twitch speed, highly-networked digital universe. They increasingly use powerful tools to play, communicate, share, support learning and solve problems. Ubiquitous hand-held technologies allow instant voice and internet connectivity as well as capture and sharing of audio and video. Word processors, spreadsheets, databases and multimedia are all commonplace in schools. Students may call upon a full range of 21st century learning technologies to research, collate and present knowledge, to design solutions and to solve problems. However, when scholastic assessment of skills and knowledge is made, access to these same tools is usually denied with *pen and paper* testing remaining the predominant mode (Curriculum Council of Western Australia, 2009). This limits the scope and form of assessment to non-performance outcomes. Constructs which cannot be tested by writing about them fail to make the test and their omission inevitably compromises the content validity of an assessment (McGaw, 2006). Students are forced to work in unfamiliar and inauthentic contexts. What assessment excludes inevitably becomes devalued and marginalised, impacting adversely on the process of teaching and learning. What is urgently required are alternative methods of assessment that will allow manageable, authentic, cost-effective, reliable and valid summative judgements of student performance; assessment methods that fit today's students and the new world in which they live and work. This study set out to investigate the use of digital technologies to support such alternative methods of assessment for an Applied Information Technology course in Western Australia. At the time, the course had only a three-hour paper-based examination for external, summative assessment.

Significance and Rationale

A critical problem exists with the form of high stakes summative assessment prevailing in the school system of Western Australia, as well as in many education systems throughout the developed world; traditional assessment fails to address the full range of learning outcomes (Lane, 2004, Lin & Dwyer, 2006). Since teachers ‘teach to the test’ (Ridgeway, McCusker & Pead, 2004) validity of assessment is inevitably compromised. This leaves curriculum authorities with a problem for as McGaw (2006) points out, “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 reduced, then those responsible for the tests bear some responsibility for that” (p. 3).

When it comes to the summative assessment of practical performance in courses such as Applied Information Technology (AIT), what is assessed typically does not match what is intended by the course. In AIT, the intention is that students will regularly use computer technology throughout the course but are not permitted to do so for their final assessment. The syllabus rationale (Curriculum Council of WA, 2009) set out in Appendix A, proposes that students should work on complex, open ended, ill-defined problems over extended time frames, devising, creating, testing and implementing digital solutions. Typically in the course, teachers attempt to present students with authentic problems to solve, leaving it to the student to choose the optimal method and most appropriate digital tools. These tasks attempt to imitate real world, problem-solving situations where there may be no single solution and no established solution algorithm. Yet the AIT final summative assessment includes none of this, largely requiring students to write what they can remember of a body of content.

In courses such as AIT, there is a further dimension to this assessment dilemma; relevance to the workplace. Innovation and problem solving skills in these domains are much sought after by employers, reflecting the increased value placed on higher order thinking skills such as analysis, synthesis and evaluation (Allen Consulting Group, 2006, p.19). Further, employers demand high levels of practical computer and information technology skills, the very skills which go unexamined in paper based summative high stakes assessment. Consequently assessment is not only misaligned with the intended curriculum but also with societal requirements.

A body of literature attests to the claim that traditional assessment methods fail to adequately assess the learning process itself and higher-order thinking skills in particular (Lin & Dwyer, 2006; Pellegrino, Chudowsky, & Glaser, 2001). The point at issue is the *validity* of summative assessment of courses such as Applied Information Technology in relation to the stated learning outcomes. It would appear self-evident that assessment of a course of study in which students learn with and through new technologies should allow students to use those technologies in the assessment process. However, at the present time, the practice exists whereby Information and Communications Technologies (ICT) are integral to learning but are disallowed during assessment. The rationale for the study is somewhat encapsulated by Ridgeway, McCusker and Pead (2004), when they state that “Skills in ICT are essential for much of modern living, and so should be a target for assessment” (p. 10).

Overview of the Study

Fundamentally, this study sought to investigate the feasibility of replacing *pen and paper* testing for high-stakes, summative assessment of Applied Information Technology with more authentic assessment tasks, with high content validity, completed and stored entirely as digital files. The quest for authentic assessment in AIT may only be satisfied by tasks which are essentially of a practical nature. Drawing on skills for the use of hardware and software, such tasks should be open-ended, production /performance measures of student ability. Two possible ways of achieving this are through a production-based, reflective process digital portfolio and a computer-based production examination and these were the two forms of assessment investigated. Eventually, the assessment tasks developed comprised a three-component portfolio and a two-component examination.

The study aimed to design, develop and implement the best assessment task possible to measure the practical performance of students in AIT. To evaluate the feasibility of this task, the study gathered data in various forms from a wide variety of sources. Qualitative data were assembled from observation and discussion with students at work on assessment sub-tasks, from discussions with teachers before, during and after school visits, from students via a survey and from teacher interview responses. Small groups of students were assembled into focus groups and responses to a series of questions were recorded and analysed. Data were also obtained from discussion with markers of students’ work. All students’ output on the assessment task was collated, collected as digital files, stored and marked and these scores, together with those awarded by the teachers, were subjected to detailed statistical analysis. Further, the study investigated the feasibility of mass storage of, and concurrent access to, digital assessment files in online repositories. Finally, the study aimed to show that the comparative-pairs assessment methodology is a practical, viable and transparent alternative to traditional analytical marking based on the "summation of the micro-judgments" (Pollitt, 2004, p. 5) made by examiners. Such a method should show no diminution in reliability and should be theoretically defensible, withstanding challenges to its credibility and eventually gaining the acceptance and public confidence currently accorded to rubric scored examinations.

Statement of the Problem

The assessment of student performance in areas such as art, science, physical education and Information Technology, does not lend itself to traditional, paper-based testing methods (see for example Hammann, Hoi Phan, Ehmer & Grimm, 2008 and Fisette et al., 2009). In these domains, much emphasis is placed on the acquisition and demonstration of practical skills and these may be difficult, if not impossible, to measure by scores on theoretical, written assessments. Alternative assessment practices, which are both valid and reliable, need to be devised for the practical aspects of these subject areas. The capture, in digital form, of students' work, may allow the development of more authentic forms of summative, high-stakes assessment with high reliability. The digital forms might be, for example, students working with application or productivity software on computer, or video recordings, audio recordings or photographs of performances, or scanned work.

Non-digital capture of students' performance has been attempted in the past with large scale assessments. For example Koretz (1998), describes some consisting of "hands-on performance tasks, especially in science in which materials or apparatus must be manipulated; hybrid group/individual assessments ... and portfolios of classroom work" (p. 313). His analysis of the quality of measurement across several states and regions in the USA revealed highly variable scoring and considerable costs in time, money and stress for these types of assessment. Conclusions such as these have dissuaded education authorities from implementing these forms of assessment in high-stakes situations. However, it is likely that the use of digital technologies to support the implementation of these alternative forms of assessment will address these barriers.

In the current assessment structure of the AIT course, the proportion of credit arising from a student's school work and the external examination is allocated equally (Curriculum Council of WA, 2009). The school mark is an aggregate of scores on various assessments which may be of the following three types (with the weighting of each type shown in parentheses): production /performance (50-60%), investigation (20-30%) or response (20-30%) as shown in Appendix A. Plainly, the course intends the majority of credit to be earned in some practical capacity. In contrast, the external examination is currently made up of multiple-choice, short answer and extended answer questions with the resulting score being used to moderate the school score. Thus up to 65% of the course score (50% from the external examination and up to 15% for school based assessment) are allocated to response. By requiring the student to demonstrate practical capability under externally monitored and regulated conditions, the balance of assessment between theory and practical might be restored to what was intended, making the assessment structure more valid and authentic.

When considering a practical assessment task, there are a number of issues which must be addressed, the first of which is the distinction between assessment of a *production* and assessment of *production and performance* (Biggs & Moore, 1993). In a production task, such as painting a picture or playing a musical piece, it is only the end product which is of interest. In contrast, a production and performance assessment, for example performing a scientific experiment, places value on both the result or end product *and* the process by which the product was developed. In AIT, the correct balance between the weightings attached to the production and process must be established.

A second consideration in the assessment of practical performance is to make the task demonstrably fair to all. The nature of the AIT course and its dependence on ICT pushes to the fore the question of equity. Schools will have different facilities, hardware and software. While standardisation of the environment for paper based assessment is simple to achieve, the same cannot be said for practical assessments involving computers and this has wide ranging implications not only for equity but also for feasibility. The nature, scope and complexity of the practical assessment task is therefore critical, as is the method of submission of the completed work; collecting in a paper is tried and tested but uploading work to a digital repository requires a lot of elements to be in place and may be difficult to guarantee to be fail-safe. High stakes assessment must be locally manageable and ultimately capable of operating in a cost effective manner over a wide jurisdiction. Any approach to assessment must garner the acceptance of stakeholders including teachers, students, parents and employers. In part, this study aimed to provide evidence to support this acceptance.

This study centred on the feasibility of replacing paper based testing of student performance, currently measured by summation of marks allocated for multiple parts, with a more holistic approach to assessment of student performance, captured and stored digitally and measured by both a marking rubric and a *comparative-pairs* method based on *Thurstone Scaling* (Thurstone, 1927). The traditional method of judging student performance is to have markers assign scores to items and to sum these to arrive at a total. An interval scale of performance is thus developed and this is used to establish an ordinal scale by assigning grade boundaries (Pollitt, 2004). Reliability is promoted by making the test items smaller and smaller so that assessor interpretations become more and more consistent leading in the extreme to some form of objective testing. However, the quest for reliability inevitably compromises validity (Ahmed & Pollitt, 2001). It is assumed that a correct macro judgement derives automatically from the summation of micro judgements but this, for reasons such as examiner question selection, context and weighting, is not necessarily the case.

What is proposed, to address the deficiencies of micro-judgements, is the repeated use of direct macro judgments in objective relative measurement of pairs of performances to establish the ordinal scale required. By repeatedly comparing performances holistically, establishing at each comparison only the better of the two performances, a scale of achievement may be developed within acceptable error limits. This method is based on the *Law of Comparative Judgments* developed by Thurstone (1927). This is a measurement model designed to establish an interval scale by making a number of paired comparisons with respect to a judge’s perceived magnitude of some property, attribute, or attitude. Trials of this method to rate student performance have indicated the potential of such a method to discriminate between traditional written examination scripts. The method is claimed to be comparable in terms of reliability, validity, quality control and cost (Elliott & Greatorex, 2002; Pollitt, 2004). The application of this method to digital representations of student performance has, with the exception of the work of Kimbell, Wheeler, Miller and Pollitt (2007), not been attempted with front-line marking. Although the theory underpinning Thurstone's work is complex, the implementation algorithm is uncomplicated and this might suggest that application of the method in this instance should be successful. However, this may not be the case and it is towards the resolution of this question that this study in part aimed to provide new evidence.

Aim of the Study

Digital technologies are already widely available in Western Australia's secondary schools. The cost of computers, video and digital cameras, peripherals and hand held devices continues to fall as their capability, functionality and availability increase. Scarcity and the expense involved in the acquisition of resources are no longer barriers to the use of new technologies in the classroom and schools offering courses such as AIT must already be adequately resourced. It is therefore not unrealistic to suggest using the same resources in the capture and submission of students' work in digital form. Particular care must be taken in the design of any AIT assessment tasks to ensure that better resourced schools are not advantaged and that access to resources is not a confounding variable in the study.

The development of alternative assessment methods, which allow the use of new technologies and are demonstrably valid, fair and comprehensive, would allow examining authorities to assess students in a realistic and educative fashion. The intention in the study was to capture students' assessment performance in digital form, for example as a data folder, a computer program or a multimedia file. These would then be uploaded to an online repository from where they would be easily and rapidly accessed by assessors and judged by both rubric based and comparative-pairs methods of marking. The combination of digital capture, online repository and comparative-pairs' judgment has been used in a pilot project by the Technology Education Research Unit at Goldsmith's College, University of London. However this was for a Design Technology course that was not specifically computer based. The results of that study (Kimbell et al., 2007), pointed positively to the feasibility of extending these methods to wider populations of students and different forms of assessment and discipline contexts.

The representation of students' summative assessment performance in digital form affords the potential for further significant advantages over traditional assessment methods. Digital files may be easily and compactly stored, transmitted, accessed and shared by markers, allowing rating of performance to be achieved by innovative methods. The capture and collation of students' digital performance on summative assessment tasks presents a further innovative opportunity which it is hoped will have far reaching cost and efficiency benefits. By making each student's performance available from an online repository, markers would have anytime, anywhere, unconstrained access to assessment materials. The combination of an online repository of digital files of students' summative performance on assessment tasks in AIT, with assessment by the method of comparative-pairs is, as far as can be ascertained, without precedent. The study aimed to add to knowledge in the use of digital technologies for the capture, storage and marking of student work. Further, it was aimed to demonstrate that alignment of the form of assessment to the stated intentions of the course better suited the preferred and required pedagogy, encouraging improved teaching and learning.

It was both timely and necessary to look at assessment practices in high stakes summative assessment of skills and knowledge in the AIT course. It was intended to investigate, from the evidence gathered in the study, the effects on the validity of assessment of new forms involving the application of digital technologies. Further, the study intended to provide a direct comparison of the reliability of assessment by holistic judgement as compared to micro-judgment methods. It may be self-evident that paper based testing of practical aspects of AIT is unsatisfactory, however, before the current approach is discarded, replacement candidates need to be thoroughly investigated and shown to be demonstrably robust, manageable and functional. It is towards this goal that this study intended to contribute new knowledge.

Statement of the Research Questions

The overarching research question for this study was:

> *How may the digital capture of students' performance most effectively support summative assessment in the senior secondary Applied Information Technology course?*

It was assumed that summative assessment should be aligned with the intended curriculum and therefore for the AIT course should focus on digital production and practical performance. The use of digital capture introduced many new factors to the assessment process. Each new aspect required an in-depth analysis to determine the extent to which it was advantageous or effectively supported assessment. Consequently, a number of subsidiary questions were addressed:

1. What are the advantages of digital capture of students' performance in support of summative assessment of practical ability in the senior secondary AIT course?
2. What are the limitations of digital capture of students' performance in support of summative assessment of practical ability in the senior secondary AIT course?
3. How feasible is the digital capture of students' performance in different forms of summative assessment in AIT with respect to
    i. Manageability,
    ii. Technical facility
    iii. Functionality, and
    iv. Pedagogy?
4. Do judgements by multiple comparisons of pairs, produce reliable scores when applied to summative assessment of practical performances in the senior secondary AIT course?
5. Which method of marking, analytical or comparative pairs, was better in assessing student practical performance in AIT?

Scope and Wider Context of the Study

This study formed one of four sections of an investigation into the feasibility of digital capture of student performance for high-stakes summative assessment in Western Australian secondary schools conducted by the Centre for Schooling and Learning Technologies at Edith Cowan University and in association with the Curriculum Council of Western Australia. Each section of the main project focussed on a different senior secondary course. The other three courses under investigation were Italian Studies, Physical Education Studies and Engineering Studies. Although this study stands alone in its findings, it also adds to knowledge in the wider context. The sharing of a common research methodology and data collection instruments allowed comparisons to be made across the four courses. Some reference is made to these in the data analysis of student attitudes towards digital assessment in Chapter Six.

Within the main project, this study formed the first of three phases and was aiming at a proof of concept, preceding a prototype and leading finally to a scalable form of assessment. The sample for this study comprised seven teachers and eight classes to trial appropriate forms of assessment, delivery and collection of materials, assessment systems, methods and reliability of scoring, with later phases refining and expanding on the findings.

Structure of Thesis

This thesis consists of seven chapters of which this, the first, has introduced the problem, presented a rationale for the study, provided an overview and listed the research questions. Chapter Two, *Review of Literature*, looks at the narrative related to the study, starting from the perspective of assessment in its broadest sense and leading on through the use of digital technologies and guidelines specific to digital assessment to the conceptual framework for the study. Chapter Three, *Method*, describes the research design, data collection and data analysis undertaken. Chapter Four, *Data Analysis*, brings together, summarises and examines the data from all sources with chapter Five, *Case Studies*, detailing on a case by case basis the data analysis, results and conclusions specific to each of the seven participating schools. Chapter Six, *Discussion of Results*, reviews the results in light of the research questions, pointing out the constraints and benefits according to the four dimensions of manageability, technical facility, functionality and pedagogy. Chapter Seven, *Conclusions*, draws out the evidence-based findings derived from the study, makes recommendations for implementation and points to some future directions in digital forms of assessment.

CHAPTER TWO

REVIEW OF LITERATURE

Introduction

This review draws upon, and brings together, two distinct and major fields of research, *Assessment of Human Performance* and *Computer Assisted Assessment*, with the aim of developing a conceptual framework and research design for the study.

Assessment is as old as learning itself. When *Homo Australopithecus* sent forth his son to kill a bear, he was unknowingly engaging in criterion referenced assessment. Had the task been set to the whole group of sons and the instruction changed to killing as many bears as possible, then the task would have become norm referenced. It is believed (Brown, 1968) that around 500 AD, the Chinese developed the first assessment system, which we understand evolved from performance based tests of martial skills to become a highly bureaucratised and selective series of written examinations leading over the course of twenty years to the higher echelons of the Imperial Civil Service. Assessment is a blanket term which describes the process of collecting data about individuals or groups for the purpose of making a decision (Biggs & Moore, 1993). Though deriving from the Latin *assidere*, to sit beside, assessment today is seldom a matter of teacher and pupil sitting side by side, negotiating the extent and quality of learning. It is closely aligned, in practice, literature and the public mind with the terms *Evaluation, Testing* and *Measurement*.

Importance of Assessment

There can be no doubt about the significance and consequence of assessment, for as Brown and Knight (1994) assert, “assessment is at the heart of student experience” (p. 11). What students regard as important, how they spend their time and how they view their academic achievements are all determined by the nature of their assessment. Changing the nature of assessment has a major impact on the scope and focus of student learning. Newble and Jaeger (1983) reported on the effects of changing the final-year assessment scheme for a group of medical students. Success in the new examinations demanded an increase in practical learning activities and this achieved the aim of realigning the practical and theoretical components to match the objectives of the course.

The critical importance of assessment in the process of learning is set out by Bransford, Brown & Cocking (2000) who claim that “assessment and feedback are crucial for helping people learn”. They suggest an alignment between assessment and learning which reflects good instructional practices, is a seamless and continuous part of instruction and provides clear feedback to “teachers, students, and parents about the levels of understanding that students are reaching” (p. 244). The importance of feedback as a positive and desired effect on learning is affirmed by Hattie (2009) in his synthesis of over 800 meta-analyses relating to achievement.

> When teachers seek, or at least are open to, feedback from students as to what students know, what they understand, where they make errors, when they have misconceptions, when they are not engaged—then teaching and learning can be synchronized and powerful. Feedback to teachers helps make learning visible. (p. 173)

Types of Assessment

The nature of learning may also be profoundly influenced by the type of assessment. The notions of deep and superficial learning derive from the seminal work of Marton and Säljö (1976). The influence of assessment on learning style was investigated by Watkins and Hattie (1985), who found that the use of closed questioning promoted superficial, reproductive learning, whereas open-ended, problem-solving tasks encouraged greater conceptual insight and deeper understanding. The use of problem centred approaches to assessment in fostering deeper learning styles is well supported in the literature, for example by the work of Thomas and Bain (1984), and Vernon and Blake (1993). For the present research, it is suggested that for the AIT course, open-ended, problem-solving tasks are well suited to delivery, completion and capture by digital means. Therefore the use of digital technologies in combination with this type of task should promote deeper learning and hence increase the validity of assessment.

This section considers two major dichotomies used to define different types of assessment. The distinction between formative and summative evaluation was first formalised by Scriven (1967). Formative evaluation aims to inform the learner of the current state of learning *during* the teaching process whereas summative evaluation aims to describe what has been learned *after* teaching is completed. Formative evaluation is continuous, diagnostic and remedial; while summative is terminal, finite and descriptive (Biggs & Moore, 1993). Formative assessment only makes sense when applied to learning objectives or criteria (so called Criterion Referenced Testing or CRT) whereas summative assessment may be based on CRT or performance in relation to other learners; so called Norm-Referenced Testing or NRT.

A second major dichotomy exists in assessment between the quantitative and qualitative traditions. Cole (1990) describes each tradition in terms of its underlying psychology, methods and values. In essence, quantitative evaluation derives from a behaviourist perspective of learning in which instruction is convergent and assimilation of content is valued. On the other hand, qualitative evaluation is situated within a constructivist view of learning, placing value on the discovery of knowledge and development of understanding from new experiences in open-ended contexts. These differences have obvious impacts on assessment method. Whereas from the quantitative position, assessment is relatively easy to carry out and favours the use of multiple choice, closed answer, right or wrong questions, the qualitative demands *authentic* assessment tasks (Wiggins, 1989), which are characterised by the demands of higher-order thinking skills and set in contexts which are as true to life and realistic as possible for the knowledge and skills under investigation.

This study was concerned exclusively with summative assessment and intended to measure a set of skills and knowledge required by the syllabus of the AIT course by examination of the products of student’s solutions to various practical and theoretical tasks.

Reliability and Validity

There are many techniques, purposes and stakeholders involved in the process of assessing learning, yet for all, the concepts of reliability and validity are paramount and interrelated (Brown, 1968). The concept of reliability, when applied to an assessment task, is analogous to the same property of other measuring instruments; we expect them to perform in an identical manner irrespective of who does the measuring and when or where the measuring is done: in other words we expect stability and consistency (Salvia & Ysseldyke, 1998). More specifically, reliability has three identifiable components which should be met; *internal*, *test-retest* and *inter-rater* reliability. Reliability with respect to generalisation to other tests, in other words that similar tests with different questions should deliver the same results, is termed *internal consistency. Test-retest* reliability describes the expectation that the same person should achieve the same score on the same test irrespective of when the tests were taken. A test must also exhibit the property of *inter-rater* (or inter-judge) *reliability,* making it be capable of being scored identically by different markers, or by the same marker on different occasions. *Unreliability* may also follow from factors external to the test itself, for example, the testing environment or the physiological state of the candidate on the day (Salvia & Ysseldyke, 1998). Estimates of test reliability are typically reported in terms of correlation coefficients which may be derived from a number of methods and subject to a variety of influencing factors (Linn & Gronlund, 1995). Some estimates of reliability are set out in Table 2.1.

Table 2.1

Methods of Estimating Reliability (Linn & Gronlund, 1995)

Method	Reliability Measure	Procedure
Test-retest	Stability	Same test, same group select a time between tests
Equivalent forms	Equivalence	Two forms of test, same group in close succession
Split-half	Internal consistency	One test. Score two equivalent halves. Correlate between halves
Inter-rater	Consistency	Two or more markers. Independently score responses

Traditionally, inter-rater reliability is achieved in high stakes summative assessment by double-marking. In this study, similar methods were used with analytical marking using a rubric being undertaken by two independent markers.

An alternative method of measuring reliability between different assessors may be derived from comparing students' work, one to another. When assessment involves *comparison* of student responses by a single examiner, the bias of the examiner is effectively eliminated. Provided that examiners decide *consistently* on the overall *relative* merit of students work, a defensible standard of reliability should be upheld and may be estimated (Pollitt, 2004). This is the basis of the comparative-pairs method of marking that was used in this study and is explained in more detail later. Pollitt suggests that reliability, in the context of high stakes summative assessment, is synonymous with precision. The quest for increased assessment precision could equally be met by collecting more and more data about students or by making assessment items less and less open to interpretation by examiners. Both of these options are undesirable; the former because it is impractical and costly, the latter because it compromises assessment validity.

The *validity* of an assessment describes its ability to measure what it sets out to measure. Validity refers to the adequacy and appropriateness of the interpretations made from assessments (Linn and Gronlund, 1995). The validity of an assessment, though simply defined, is, like reliability, an equally complex and multi-facetted concept invoking many kinds of evidence (Messick, 1994). Validity is not a property of an assessment but of the specific inferences which may be drawn from the results of the assessment.

Messick (1996) views validity as an integrated concept made up of six clear and interdependent aspects which must not be viewed in isolation but as complimentary forms of validity evidence. *Content validity* refers to the features of the domain under test that the assessment intends to reveal. Increasing levels of performance must reflect increased complexity of knowledge and skills directly relevant to the construct under examination. The *Substantive validity* is concerned with the suitability of the sampling and coverage of the content under review. *Structural validity* describes the consistency of the assessment and scoring process and *Generalizability* describes the extent to which other tasks might equally represent the construct or aspects of the construct. *External* aspects of validity describe the degree to which other behaviour and performance measures are consistent with the score generated by the assessment. *Consequential* aspects of validity describe the uses, interpretations and implications, both intended and unintended, resulting from the assessment score. Validity, therefore, may have many and various aspects and some of these are described in more detail in Table 2.2.

Table 2.2

Some Major Properties of Validity (adapted from Salvia & Ysseldyke, 1998)

Property	Method	Meaning	Example
Face	Experts examine the test and seek to agree on the scope, relevance and intention.	Does the test appear to be measuring what it intends to measure?	A panel of expert judges analyse a test and then either allow or reject certain questions
Content	Comparison of the assessment task is made to the domain to be assessed	How precisely and comprehensively does the sample of tasks in the assessment represent the domain to be measured?	A student is asked to learn the definitions of 100 words and then tested on a sample of 10
Predictive	Comparisons of future performances are made with the current assessment task	Does knowledge of a person's current score portend accurately to scores on related tests at some future time?	Aptitude tests of coordination undertaken before air pilot training.
Consequential	A reasoned analysis of the consequences of a test for teaching and learning is undertaken	What are the impacts of the test on the wider interpretations, uses and consequences beyond its result?	The use of multiple choice questions may promote superficial learning styles
Construct	Logical inferences are drawn from a variety of data types	How well do the test data measure the construct under investigation?	Have students 'think out loud' as they perform tests of verbal reasoning.

Salvia & Ysseldyke, (1998) propose that "Validity refers to the appropriateness, meaningfulness and usefulness of the specific inferences that can be made on the basis of observations or test results" (p. 166). How we observe and measure performance, the method of task assessment, depends largely on the nature and purpose of the assessment task. The concept of *Validity* was at the core of this study and formed the key concept under investigation. The purpose of the study was to find ways of improving the validity of assessment in the AIT course whilst at the same time maintaining or improving reliability. It could be argued that paper based assessments of the AIT course have poor validity in all aspects except perhaps the predictive: success in one paper based examination may be a good indication of potential success in another. However, digital forms of assessment might be expected to show high validity with regard to the all the general validity criteria of performance assessment.

Purposes of assessment

Each assessment is conducted for its own specific purpose and this will influence not only the type of assessment but also the relative meaning and importance of its validity and reliability. Brown (1997) lists the purposes of assessment from the perspective of students, society and teachers, and a summary of these is presented in Table 2.3. Each purpose suggests a process and method; CRT or NRT, qualitative or quantitative. For example, results of school leaving examinations are quantitative, norm-referenced scores or rankings which may be used as a license to proceed to tertiary courses. Informal self-assessment may be used to inform a student of his or her current progress and are criterion referenced and qualitative, perhaps providing motivation for further study. The purpose of the assessment investigated in this study was the ranking of candidates used as the basis for entrance to tertiary institutions.

Table 2.3

Purposes of Assessment (adapted from Brown 1997, p. 11)

Student	Society	Teacher
feedback	Pass or fail	Feedback
motivation	Grade or rank	Improvement of teaching
diagnosis	License to proceed	Course evaluation
self-assessment	Select for future study	Quality assurance
profile	License to practice	
	Predict success in future	
	Employment selection	

Performance Assessment

A second basis of classification of assessment type, as opposed to the formative and the summative divide discussed previously, is the division between fixed choice and complex-performance assessment. Fixed or multiple choice assessment, requires respondents to select the best possible answer from a list of options. Examples of complex-performance assessment include “open-ended problems, essays, hands-on science problems, computer simulations of real world problems, and portfolios of student work” (Linn, Baker & Dunbar, 1990, p. 2). The relative strengths and weaknesses of each form are shown in Table 2.4.

Table 2.4

Comparison of Fixed Choice and Complex-performance Assessment (From Resnick & Resnick, 1995)

	Fixed Choice Test	Performance Assessment
Learning Outcomes	Efficient measurement of factual recall. With skilled question writing can measure higher order thinking. Inefficient for select/ organise, writing and some problem solving skills.	Inefficient measurement of facts. Suitable for measuring understanding, thinking skills, originality, correspondence with learning objectives.
Question preparation	Large number of questions needed. Good questions take time to create.	Small number of questions required.
Course content	Large number of questions means good sampling.	Small number of questions means sampling is limited.
Student response	Highly controlled. Avoids irrelevant variance (e.g. effect of handwriting) Guessing possible.	Free response fosters creativity and originality.
Scoring	Objective.	Subjective.
Effect on learning	Promotes comprehensive knowledge. Can promote higher order skills if properly constructed.	Promotes organisation, integration and expression of ideas and concepts.
Reliability	High, may be computer scored.	Low due to inconsistent marking.

Critics of fixed choice tests, for example Resnick and Resnick (1992), report the tendency to over emphasise lower level skills, such as factual recall, at the expense of the higher order skills of problem solving. Further, this type of testing steers teaching and learning along a path which is in contradiction to currently accepted theories of learning, by emphasising and rewarding the accumulation of unrelated facts and skills. The belief that learners actively construct knowledge, based on the interplay between new and previous experience in social contexts, supports the use of complex performance assessments. Extended tasks such as essays, laboratory experiments and oral presentations, are better and more closely aligned to the instructional goals of teaching students to think and solve problems (Darling-Hammond & Anderson, 2010, p. 7). Complex-performance assessments are often referred to as realistic problems or authentic tasks, reflecting the intention to have students solve real world problems in true-to-life contexts. Though assessment of complex performance tends to be more time consuming than fixed choice testing, requiring training and subject area expertise, it is generally accepted that both forms of assessment have a role to play.

The *type* of performance assessment may be described and classified in a variety of ways. Cronbach (1990) places assessment into two broad categories which are distinguished in terms of the use to which the results are put. Measurement of maximum performance, describes the case in which a candidate is motivated to achieve as high a score as possible, for example in aptitude or achievement testing. Alternatively, in measurement of typical performance, the concern is more with obtaining representative responses, for example in assessment of attitudes and personality traits. This study was concerned with the assessment of maximum performance.

Assessment of Practical Performance

The meaning of performance assessment is by no means clear in the literature being closely allied with the terms product, process and authenticity (Palm, 2008). Fitzpatrick and Morrison (1971) noted that performance assessment was synonymous with performance-and-product assessment. Messick (1994) points out that the call to investigate performance and product assessment may be traced back to the 1960s with an “upsurge of renewed interest” in the late 1990s which had “positive consequences for teaching and learning” He argues for the need to address “issues of validity, reliability, comparability and fairness” (p. 13). These are social values which may be the intended or unintended consequences of the assessment emanating from considerations of its purpose and domain. He argues that in subject areas such as the performing arts, the product and the performance are one and the same thing, for example in the assessment of proficiency with a musical instrument or of acting skill. In others, such as painting a picture, the diversity of possible techniques makes assessment of process meaningless and it is only the end product that counts. In cases such as these, assessment makes no inference as to the underlying skills and knowledge of the student. In other subject areas, such as scientific experiment, both the end product and the process are important since correct procedures, for example safety practices, are also of value and are amenable to assessment.

Despite the compelling arguments for performance assessment, such methods have found only limited application. Lane (2004) believes that a decline in the use of performance assessments in the USA derives from increased accountability and constraints on resources. This in turn has led to a misalignment between assessment and instructional practices to the detriment of eliciting higher order thinking. Colley (2008), points out that although performance based assessment is soundly based in cognitive theory, there are limitations, for example the additional time involved in preparation, documentation, testing and gathering resources. A further issue is the assessment process itself. “Performance tasks vary greatly in terms of the subject matter addressed... it can be difficult to evaluate student performance...at district or state level the costs of administering and scoring are three to five times higher than those of conventional testing methods” (p. 70).

Zane (2009) sets out three underlying principles for the development of performance tasks. Firstly, developers should “identify contextual components of the task” such as the cultural setting, situation and tools available. Secondly “determine the nature of the problem and how ill-structured the task should be” (p. 87) bearing in mind that too little structure makes for increased difficulty of measurement and interpretation of scores. Thirdly, the activities allowed and permitted modes of response should be determined, with task parameters such as time allowed being specified.

> In the real world, competent practitioners resolve problems by gathering information, considering pros and cons of situations, using social interaction and/or collective problem solving, confronting ineffective strategies as needed, exploring misconceptions, making decisions, self-evaluating, and self-correcting mistakes. Exactly how the competent practitioner responds to the problem may include taking certain actions, writing a report, or responding in some other way. (p. 88)

In this study, the responses constituted two distinct forms; portfolio and examination. Each of these has its own constraints and benefits.

Koretz (1998) analysed four large scale portfolio assessment systems in USA schools and reported problems with resources, reliability and manageability. Each study involved assessing students for comparability across schools and across states on the basis of performance in Mathematics and English. Koretz concluded that there was insufficient “evidence that the resulting scores provide a valid basis for the specific inferences users base on them” (p. 333). Assessment by portfolio, though desirable and important was perceived as fundamentally difficult.

Crawford and Fakete (2006) assert that there is often a mismatch between the intention of a practical assessment and the knowledge and skills actually assessed. The target of the examiner may be to assess a student's understanding "as distinct from rote-learning, speed or other aspects of their intellect" (p. 185) but an analysis of student results on 16 examination questions in a second year Computer Organisation course revealed that several were tackled by students in ways quite different from what the instructor had expected.

In summary, assessment of practical performance has many associated issues and complications; it may be difficult to create, administer and score as well as being time consuming and costly. The use of Computer Assisted Assessment may afford opportunities to alleviate some of these aspects.

Computer Assisted Assessment

The use of ICT in assessment is not new, having its origins in the optical mark recognition (OMR) of the 19th and 20th centuries (Bull, 2004). The term Computer Assisted Assessment (CAA) is now used to describe any form of assessment of student learning in which computers are used, subsuming earlier but still current terms such as Computer Aided Learning (CAL), Computer Adaptive Testing (CAT) and Computer-Based Assessment (CBA). Though exclusively linked in the past with multiple choice test formats, CAA in the 21st century affords the potential to expand assessment practice to include assessment tasks based on multimedia, simulations and virtual worlds. A variety of reasons have been suggested for using CAA and some advantages and disadvantages are summarised in Table 2.5.

As discussed earlier in this chapter, the strong link between what is assessed and what is learned is well supported in the literature (see for example Beevers, Foster, McGuire & Renshaw, 1992). Students may be reluctant to invest time in any activity which does not directly impact upon their final grade. Without increasing the total time spent on assessment, the efficiencies afforded by ICT may allow the frequency and scope of assessment to be increased, thereby expanding the range of knowledge assessed and promoting study of all parts of a course. Increasing assessment frequency promotes student practice and revision allowing improvement in feedback to both learners and teachers.

Table 2.5

Advantages and Disadvantages of CAA (Bull 2004)

Advantages	Disadvantages
Progress monitored through frequent assessment	Initial cost and time overheads may be large
Promotion of student awareness and self-assessment	Hardware failure during high-stakes testing
Detailed feedback to students	Student IT skills must be present in advance of assessment
Students acquire IT skills	Requires training of assessors and invigilators
Increased assessment frequency is made possible by automated delivery and marking	Requires coordination between teachers, learners, IT staff
Scalable to large groups of students	Limited question types
Incorporation of multimedia in assessments	Requires high skilled and time consuming question development if higher order skills are to be tested.
Advanced statistical analysis of questions and responses made easy	
Diagnostic assessment facilitated	
Administrative efficiencies in entering and transmission of marks	
Double marking made redundant	
Human error eliminated	
Assessment on demand	
Adaptive assessment, based on student responses, made possible	
Randomised selection of questions and distracters aids security	
Question bank sharing	

CAA may also extend the range of assessment methods available, negating over-reliance on any particular mode. For example, the inclusion of multimedia offers the potential for innovative and varied assessment tasks beyond the range of pencil and paper assessment. This might be an important motivational aspect, which together with timely and informative feedback has been shown to have benefits for the overall performance of students on a course (Schmidt, 1990). The related issues of consistency of assessment scoring over an ever increasing volume of assessments and the possible need to defend an individual student's course or examination mark against a challenge, favour the use of objective testing using CAA (Knight, 2002). Further, efficiencies in assessment administration and reduction in marking loads are important factors favouring the adoption of CAA.

e-Assessment

A definition of e-Assessment is "the end-to-end electronic assessment process where ICT is used for the presentation of assessment activity, and the recording of responses...from the perspective of learners, tutors, learning establishments, awarding bodies and the general public" (JISC, 2007, p. 6). The ability to deliver and capture student assessment performance in digital form has many potential advantageous implications. These range from "doing traditional things in new ways, to extending what we could traditionally do, and onwards to supporting learning in new ways" (BECTA, 2006, p. 3). The progression from supporting summative assessment to supporting the learning process is made evident in Table 2.6.

Table 2.6

Levels of Functionality of e-assessment (from BECTA's View, January 2006)

The scanning of examination scripts for electronic forwarding to markers, and online marking Electronic delivery of examination scripts, printed out at the examination centre Online delivery and completion of 'traditional' examination scripts (paper behind glass)	Improving traditional processes through technology
Online delivery of tests and examinations which utilise technology to extend what can be assessed, for example the use of multimedia, simulations and 'drag and drop' mechanisms The extension of assessment, with the availability of item banks and randomised question choice, to provide assessment on demand	Extending the limits of traditional practice through technology
Use of e-assessment to provide on-going formative assessment (with assessment integrated with learning content), progress tracking, goal setting and feedback to the learner and practitioner Use of e-assessment to diagnose understanding and levels of ability before a course of study is undertaken; this is already in place in some sectors with key and basic skills tests Use of e-portfolios to enable the recording of achievement and storing of evidence for a longer period and for more varied purposes than an examination, and to store evidence of varying types (for example audio and video files of practical work)	Technology in the service of learning

Used imaginatively, e-assessment may allow the scope and nature of what is assessed to be extended and improved (BECTA, 2006). Assessment can be made more appropriate, that is, a closer alignment between what is taught and what is assessed may be forged. Assessment can become on-demand and potentially scalable to large numbers of candidates. Assessment experiences may be made more *authentic* through the use of real-world problem scenarios, perhaps set in virtual worlds, requiring original solutions involving a range of software and hardware. When objective forms of testing are seen to be suitable, delivery of questions which are interactive and multimedia may promote student engagement; computer marking and analysis may endorse claims of increased reliability and validity.

This study was concerned with a sub-set of the e-assessment spectrum; the use of assessment tasks to generate digital files or *Digital Forms of Assessment*. From an administrative viewpoint, digital forms of assessment allow easy collation, storage, back-up and transmission of student performances, savings in printing, paper and transportation costs and improved efficiency in scoring and storage of marks (BECTA, 2006). Historic records of student performance, gathered over many years, can be used as evidence in support of maintained standards and quality assurance.

> There is no doubt that the use of computer technologies brings with it an opportunity to revolutionize the delivery and assessment of learning outcomes. It has the potential to make fundamental changes in how we teach; which mental processes, skills and understandings we measure; and how we make decisions about student learning. (Taylor, 2005, p. 11)

e-Assessment of Complex and Ill structured Tasks

Increasing the authenticity of assessment presents challenges for measurement of student performance. If assessment tasks are to be grounded in real world situations, they are likely to be more complicated and time consuming than abstract and idealised assessments tasks which commonly make up paper-based tests. How might it be possible to assess, reliably and validly, student performance in these circumstances? Spector (2006) suggests that the difficulty of assessment is one of the reasons "such problems are often avoided in school-based instruction…A persistent problem with regard to evaluating problem-centred approaches to learning is that there is not a well-established and reliable methodology to determine learning outcomes" (p. 111).

One approach to this predicament has been attempted in the domain of Design and Technology with the *e-Scape Portfolio Assessment Project* (Kimbell, Wheeler, Miller & Pollitt, 2007). The project centres on the creation, in real-time and in digital form, of a student portfolio during the completion of an extended design assessment task for the purposes of summative assessment. The exact nature and form of the portfolio are described by Kimbell et al. as:

> ...neither a container nor a reported story, but rather a dialogue. The designer is having a conversation with him/herself through the medium…So it has ideas that pop up but may appear to go nowhere- and it has good ideas that emerge from somewhere and grow into part solutions- and it has thoughts arising from others comments and reflections on any ideas...It is more like a designers sketchbook – full of notes and jotting, sketches, ideas, thoughts, images, recordings and clippings. (p. 8)

The portfolio was captured entirely in digital form using a Personal Digital Assistant onto which sketches were drawn and digitised, audio conversations and comments were recorded and photographs were taken and stored before being periodically backed up to a central server. The *e-Scape Portfolio Assessment Project* set out to test the feasibility of an assessment method which was neither a formal examination nor a piece of coursework but an extended production examination. Students assembled a digital portfolio over a period of six hours on two days, according to a predetermined template of activities. The portfolio consisted of drawings, photographs, voice memos and notes, created on a hand held computer and uploaded to a custom-built, web-based database. From the first page of this, which consisted of 22 thumbnail screens, the students work was able to be reviewed in detail and assessed.

The marking of student performance was also radically different in using a *comparative-pairs* method advocated by Pollitt (2004). He states that the examination system requires examiners to,

> ...assign to each of a large number of students a number which represents their level of performance on tasks which are designed to discover their level of educational achievement in some educational area….to sort the candidates into a rank order with sufficient precision and categorisation to meet the needs that our national educational, economic and political systems place on the examination system, and to attach constant standards to that ordering. (p. 3)

In short, the purpose of summative assessment is to *judge* the overall quality of students on a standard ordinal scale, and scoring and totalling the marks awarded to several micro-judgments, is just one way to achieve this. Current methods of summative assessment, with their focus on summation of micro-judgements are “dangerous and ...several harmful consequences are likely to follow” (p. 5) with accurate and consistent measurement of a student’s ability unlikely to be achieved.

As an alternative, Pollitt (2004) suggests the method of paired, comparative, holistic judgment. The essence of the method is the comparison of two examination scripts, the outcome being a decision as to which of the two has more merit. Further judgments of the same and other papers are made until the relative order of merit of all the papers is established and an ordinal scale of achievement created within acceptable error limits. In the process, each script would need to be available for comparison on demand and this is made possible by digital and communications technologies. By extension, Kimbell et al. (2007) have demonstrated that making similar comparisons of digital portfolios is both feasible and desirable, producing a scale of assessment results with high reliability.

Phase 2 of the *e-Scape Portfolio Assessment Project* concluded by addressing findings in relation to four research strands: *technological*, *pedagogic*, *manageability* and *functionality*. With respect to the first of these, the system met the technological challenges and was "sufficiently robust to be taken to 14 schools during the national pilot in which 300 learners undertook studio/workshop activities and successfully uploaded their portfolios into the website" (Kimbell et al., op cit, p. 95). Further, from a pedagogic viewpoint, the report concludes that learners responded favourably and were sufficiently engaged by the structure and activities. The forced compliance with the progressive structure, from design ideas through to finished product, provided support to learners encouraging and promoting them to produce their best work.

Manageability issues centred on the use of hand-held devices in the rough and ready environment of the workshop. The report concluded that "the approach adopted for e-scape was indeed manageable for learners, for teachers and for the research team" (Kimbell et al., op cit p. 96). With regard to functionality, the e-scape system was found to offer a workable and innovative solution to the assessment of performance tasks which are notoriously difficult to manage equitably and to assess reliably. The new assessment model of comparative-pairs linked to a web based portfolio produced reliable statistics which attest to the overall functionality of the system.

Assessing Students on their Application of IT Skills and Knowledge

Kennedy (2008) defines digital literacy as "the ability to find, interpret, comprehend, understand, evaluate, restructure and re-purpose the wide variety of media types that can be stored, retrieved and manipulated using a computer" (p. 228). The AIT programme is in essence a course in digital literacy and this definition concisely summarises both its intent and nature. AIT outcomes stipulate assessment of both the product and the processes used in the development of information solutions, making the assessment of practical ability obligatory (Curriculum Council of WA, 2009). Indeed the focus of the course is the *practical* use of ICT in the generation and communication of solutions to authentic, real-life problems.

There are a number of ways of providing a practical assessment component to technology courses. In the USA, Educational Testing Services (ETS), the creator of the Scholastic Aptitude Test, developed an ICT Literacy Assessment which aimed to measure the ability to use technology as a tool to research organize, evaluate and communicate information (ETS, 2002, p. 17). The developers initially targeted post-secondary students with web delivered scenarios presenting test-takers with a series of 16 simulated tasks such as advanced searching, sorting, organising, presenting and communicating information. The report differentiates between tasks designed to assess proficiency, in which assessment of the product is holistic, and tasks designed to assess and diagnose skills in ICT components, namely the accessing, management, integration, evaluation and creation of information solutions.

In the UK, the National Assessment Agency has developed an 'on screen' Key Stage 3 (11-14 year olds) ICT test which became statutory in 2008 (Qualifications Curriculum Authority, 2008). Extensive piloting and trialling was undertaken in 2006 and 2007. Schools wishing to take part in the trials were required to be audited and approved as testing centres ensuring that they had the technical capability to manage and deliver the test. Planning and preparation of staff and pupils was provided and included practice tests. Feedback indicated a wide disparity in the time required to organise the event and the test software itself presented a few problems when used in conjunction with some workstation configurations. Activities in the test included following hyperlinks, adding information to databases, identifying mail merge fields, correcting errors in spreadsheets, formatting and saving. The test, like the trials, was automatically marked with the evidence gathered every time an answer was attempted being made available at the closure of the test in the form of a report.

Design and Development of Digital Assessments

Romeo (2008) sets out some basic principles for the effective design of learning environments suggesting that problem based or project based learning should be the “dominant instructional strategy” with students employing “technology to research solutions and present the product of their investigations” (p. 214). Further, assessment should be authentic, promoting and emulating the goals of learning and providing opportunity for learner reflection. In essence Romeo calls for tasks which not only engage the learner but also build new knowledge. Whether assessment will do as Romeo suggests will depend on the design of the assessment. Therefore, in recent times, authorities and researchers in many parts of the world have developed guidelines for the use of digital technologies in the assessment process.

The Qualifications and Curriculum Authority of UK sets out 13 regulatory principles for all forms of e-Assessment (Qualifications and Curriculum Authority, 2007). The principle of fairness, comprising the aspects of validity and reliability forms the first of these. “Awarding bodies must ensure that assessment delivered and maintained by electronic means is fit for purpose and produces a valid and reliable measure of a candidate’s skills, knowledge, understanding and/or competence”. (p. 7) Principles six and seven are concerned with fairness in a different sense, specifically that access to assessment should not be limited by physical disability, by ensuring that “disabled learners are not treated less favourably than non-disabled learners” (p. 11) and that “the use of technology does not create barriers for learners…by enabling familiarisation and/or training sessions appropriate to the mode of delivery”. (p. 11)

The British Psychological Society (2002) published a set of general guidelines for Computer-Based Assessments through its Psychological Testing Centre. These guidelines include sections on the use of digital technologies in *Assessment Generation, Assessment Delivery, Assessment Scoring* and *Interpretation, Storage, Retrieval* and *Transmission*. The guidelines are defined from the perspective of assessment developers and users. In a further example, the Council of the International Test Commission (2005) have developed international guidelines for good practice in computer-based and Internet delivered testing. These focus on four issues: technology selection, quality of the testing, the test environment and testing security. All contexts under consideration 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.

Guidelines Specific to Computer-Based Examinations

Computer-based examinations involve students sitting at computer workstations completing tasks, including typing answers to questions. They may be required to use various pieces of software to create digital products or simply respond to questions delivered via a testing engine through a browser. In AIT, while both types of assessment activities could be involved, it is likely, given the nature of the subject matter, that the focus would be on creating products in digital form. The key issues for computer-based examinations are set out in comprehensive fashion by Parshall, Spray, Kalohn and Davey (2002). Though principally concerned with administration, scoring and analysis of low-complexity item types the authors offer comment on computer-based examinations involving innovative item types with high complexity. Generally, such assessments are composed of fewer items and require greater time allocation but “provide a great deal more information than would be available from a single, discrete multiple-choice item”. The reduced number of items,

> ...can result in problems of limited task specificity and poor generalizability for the assessment as a whole…and may be especially subject to memorability and item exposure. However the related advantages for extensive, complex and integrated tasks also hold, in that they can provide a more contextualised assessment and a better real-world congruence. (Parshall et al, 2002, p. 84)

The International Test Commission has provided detailed guidelines for computer-based examinations (The Council of the International Test Commission, 2005). These guidelines are specific to test developers, test publishers and users and mainly relate to response type assessments. An array of specific guidelines is suggested as set out below.

1. Give due regard to technological issues in Computer-based Testing (CBT) and Internet testing
    - i) Give consideration to hardware and software requirements
    - ii) Take account of the robustness of the CBT/Internet test
    - iii) Consider human factor issues in the presentation of material via computer or Internet
    - iv) Consider reasonable adjustments to the technical features of the test for

candidates with disabilities

   v) Provide help, information, and practice items within the CBT/Internet test

2. Attend to quality issues in CBT and Internet testing
   i) Ensure knowledge, competence and appropriate use of CBT/Internet testing
   ii) Consider the psychometric qualities of the CBT/Internet test
   iii) Where the CBT/Internet test has been developed from a paper and pencil version, ensure that there is evidence of equivalence
   iv) Score and analyse CBT/Internet testing results accurately
   v) Interpret results appropriately and provide appropriate feedback
   vi) Consider equality of access for all groups
3. Provide appropriate levels of control over CBT and Internet testing
   i) Detail the level of control over the test conditions
   ii) Detail the appropriate control over the supervision of the testing
   iii) Give due consideration to controlling prior practice and item exposure
   iv) Give consideration to control over test-takers authenticity and cheating
4. Make appropriate provision for security and safeguarding privacy in CBT and Internet testing
   i) Take account of the security of test materials
   ii) Consider the security of test-taker data transferred over the Internet
   iii) Maintain the confidentiality of test-taker results

Many of the guidelines apply generally to any test-taking context, whether computer-based or not. Many of the other guidelines were not applicable to the current study, for example those in 4i to 4iii, because only single classes and their teachers in particular schools were involved. While mainly relevant to the implementation of large scale online testing, many of the guidelines in areas one to three were relevant to the study. In essence, the first three sets of guidelines were addressed by the *Feasibility Framework* used in the study, aligning with the *Technical*, *Functional* and *Manageability* dimensions as set out in the *Definition of Terms* p. xix.

Recent Implementations of Computer-Based Examinations

In the last five years, there has been increased international interest in computer-based testing to assess ICT capability. For example, an international research project, the *Assessment and Teaching of 21st Century Skills Project* has commenced, supported by the three computer companies *Cisco*, *Intel* and *Microsoft*. There have also been trials of such tests in a number of countries including the UK, Norway, Denmark, USA and Australia (MCEETYA, 2005). The trial in the UK involved a multi-million pound simulated system accessed by students through their school computers. In the Norwegian example students used their own government-provided notebook computers. In the USA a decision has been made to include an ICT literacy test in national testing in 2012 but in a number of states such tests are already in existence. In Australia, the Australian Council for Educational Research used a computer-based test to assess the ICT literacy of Year 6 and 10 students (MCEETYA, 2005). They developed the test around a simulated ICT environment and implemented the test using sets of networked laptop computers. While they successfully implemented the test with over 7000 students, this was over a long period of time. The use of a simulated environment for an AIT examination would be expensive to develop, difficult to scale up and unable to provide a great enough variety of activities for year on year implementation. A simpler and more cost effective approach has been trialled by Fluck, Pullen and Harper (2009) and involves the delivery of the whole examination environment on a single CD ROM or flash drive. Based on free open-source software and designed to run on students' own laptop computers, the system allows a secure examination to be conducted and invigilated without the need for specialist ICT skills.

Guidelines Specific to Digital Portfolios

Barrett (2005) defines a digital portfolio "as a container, allowing students/teachers to collect and organise portfolio artefacts in many media types (audio, video, graphics, text); and using hypertext links to organise the material, connecting evidence to appropriate outcomes, goals or standards" (p. 5). 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 requires a precise structure with limits on the type and size of the contents, control on the time available and authentication of the work by the teacher and the student. All of these features were incorporated into the portfolio assessment used in the study.

In their review of e-assessment, Ridgway et al. (2004) promote the use of digital portfolios as a means of overcoming the ‘distinction’ between academic and practical subjects stating that:

> ...abstract thinking is important; appropriate action in context that rests on practical competence is important. Neither is much use on its own, and students should be taught to both abstract and apply. For this to become a classroom reality, assessment systems must require students to show the full spectrum of competencies in a number of school subjects. If high-stakes assessment systems fail to reward such behaviours, they are unlikely to be the focus of much work in school. E-portfolios offer a way forward. (p. 26)

Carney (2004) developed a set of critical dimensions of variation for digital portfolios:

i) Purpose(s) of the portfolio;
ii) Control (who determines what goes into the portfolio and the degree to which this is specified);
iii) Mode of presentation (portfolio organisation and format; the technology chosen for authoring);
iv) Social Interaction (the nature and quality of the social interaction throughout the portfolio process);
v) Involvement by the teacher. When considered more broadly, other important portfolio participants might include other students and parents.
vi) Use (can range from low-stakes celebration to high-stakes assessment).

However, because in the present study the purpose was high-stakes assessment, this reduced the potential for variation.

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

The portfolio developed in the present study, being prescribed by the research team, included all of these characteristics.

Beetham (2008) found that e-portfolios are “less intimidating for some learners than a traditional examination” and “provide evidence that gives a much richer picture of learners’ strengths and achievements than, for example, a test score” (p. 4). She points to the need for web-based relational database systems to implement portfolios. Whilst in the past, e-portfolios have been found to take longer to moderate and mark, this has become more streamlined when part of an “integrated assessment facility”. She provided five commercial examples of such systems listing “issues relating to the use of e-portfolios for summative assessment” (p. 5). Seven of the nine issues are technical and most are addressed by the use of a good assessment management system. The remaining issues are:

- Acceptability and credibility of data authenticated by Awarding Bodies
- Designing assessment strategies to make effective use of the new tools and systems
- Ensuring enhanced outcomes for learners, e.g. higher motivation, greater choice over evidence, assessment around capabilities and strengths

Although the present study did not use an on-line system due to school network constraints, the marking did use an on-line database accessed through a browser.

She also raises some issues for teachers and learners (p. 16). These are the fit with existing practices and expectations, degree of access to and ICT capability of teachers and learners and acceptability and appropriateness of e-portfolio use. These issues were critical to the present study and were investigated by the collection of qualitative data.

Digital portfolios clearly offer scope for summative assessment and this study hopes to add new knowledge to the practical implementation considerations in the specific area of Applied Information Technology.

Conceptual Framework for the Study

Drawing from the ideas discussed in the preceding review of the literature, a conceptual framework was developed to underpin the study. The key concepts and relationships are represented in the diagram shown in Figure 2.1.

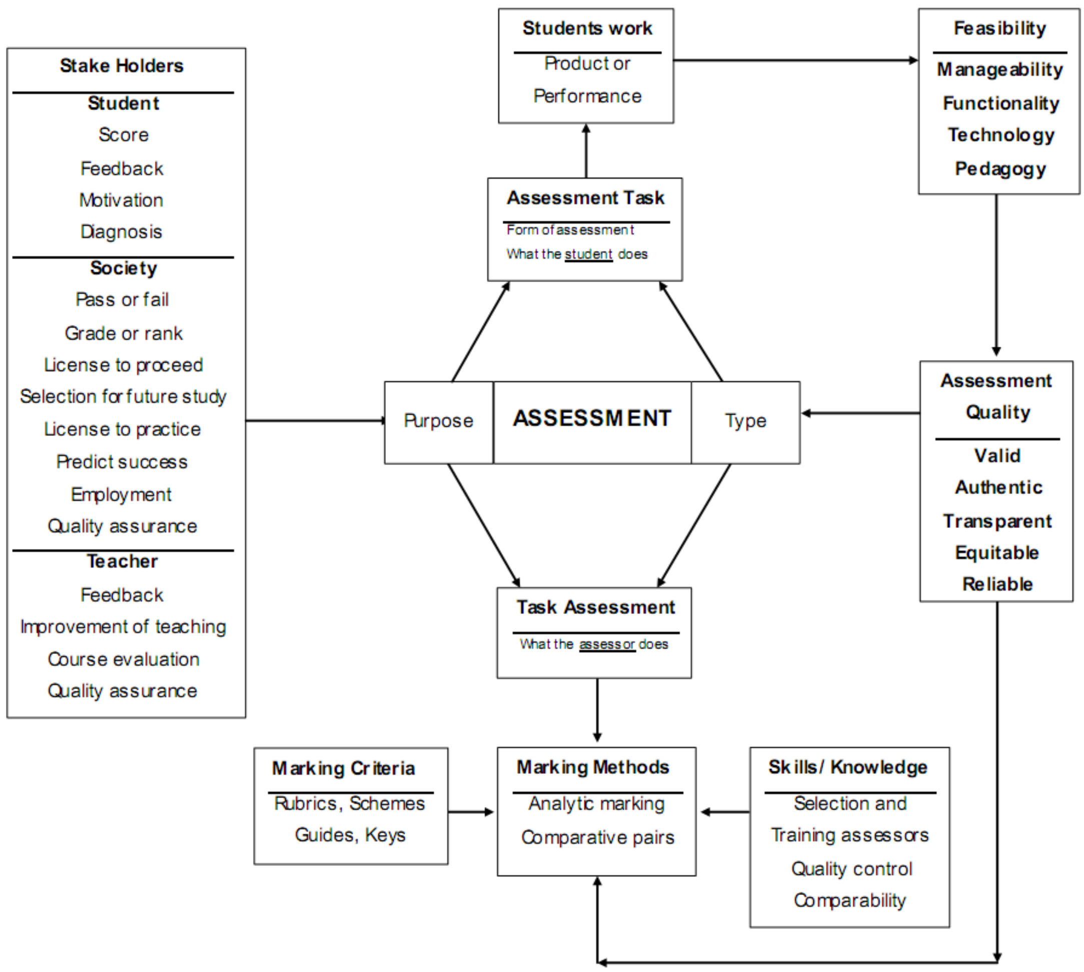


Figure 2.1 The Assessment process: composite from Brown (1997), Kimbell (2007) and Campbell (2008)

In the discussion which follows, terms appearing in the diagram are highlighted using italics. The conceptual framework draws on the work of Campbell (2008) who makes clear the distinction between the *assessment task*, what the student does, and the *task assessment*, what is done by the assessor. Central to the study was the concept of *assessment* of student performance. An assessment has a *purpose* and is of a particular *type*. The *purpose* of the assessment is critical to all aspects of the design and implementation of the *assessment task* and to the process of *task assessment*. The *type* of assessment should meet the *assessment quality* guidelines and must be amenable to *reliable* marking.

The assessment *Type* encompasses the formative/summative and qualitative/quantitative divides described previously in this chapter. Assessment *type* might be formative/qualitative, formative/quantitative, summative/qualitative or, the type of assessment currently employed in the AIT course, summative/quantitative. This *type* of assessment however, fails to meet the *assessment quality* criteria of *validity* and *authenticity*, placing undue emphasis on the assimilation and reproduction of content to the detriment of higher order thinking, creativity and understanding set in contexts which are as true to life as possible. For these reasons, summative/qualitative assessment was the chosen *type* and the starting point of an iterative cycle of task development, feasibility and quality as shown in Figure 2.1. The *forms* of assessment deemed best suited to summative/qualitative assessment *type* were a digital portfolio and a computer-based, performance examination.

Assessment Q*uality* refers not only to the *reliability* of the marking process, for whatever the *type* of assessment chosen, it must be amenable to reliable scoring, but also to the general fairness and fitness for purpose of the *assessment task*. These properties of *assessment quality* are described by validity, authenticity, transparency and equity.

Both the *purpose* and the *type* of assessment determine what is required of the student, in other words the *form* of assessment. The *purpose* of assessment is determined by the *stakeholders*. Since the aim of this study was to examine an alternative to the current approach (the summative assessment of students' ability in AIT) the main *purpose* of assessment remained unchanged; to deliver a *score* to each student indicative of their ability in the AIT Stage 2 course. The meaning attached to that score is inextricably linked to the perspectives of other *stake-holders* in the assessment process, principally those listed under *Society*. The *score* would contribute to a *rank* forming the basis of a *license to proceed* to *future study*. It might also have predictive validity for *future success* in this subject area or be a selection factor in *employment*. *Teachers* are also important stakeholders but for them the *purpose* of assessment is different; to provide feedback on their teaching, leading to *evaluation* of method and possible *improvement*. For *students*, assessment is a primary factor in *motivation* to study with the score awarded providing *feedback*, *diagnosis* and *motivation* towards further study.

The *purpose* of assessment also has implications for the *type* of assessment task. Students and teachers require the assessment task to be of a fair level of difficulty and matched to the course outcomes; society requires that the level and scope of the assessment task meet *quality assurance* guidelines. In this study, the syllabus for AIT set out the scope and level of possible student activities and these in turn set the degree of difficulty of the assessment tasks.

As well as matching the *purpose*, the *type* of assessment had to be of demonstrably suitable *assessment quality*. Principally, the *type* had to be *transparent* in its requirements and *valid* and *authentic* in terms of the demands made on students. In this study, the drive for *assessment quality* determined the assessment *type* and shaped the *assessment task*, requiring students to create a *product* in an extended portfolio or to demonstrate skills with some form of practical *performance* examination. The *type* of assessment employed, had to meet *feasibility* criteria whilst at the same time allowing students to demonstrate their skills and knowledge to the required breadth and depth as set out in the course syllabus.

Feasibility of the assessment *task* was judged in terms of a *feasibility* framework, adapted from Kimbell et al. (2007), and consisting of four dimensions; the *technical*, the *manageable*, the *functional* and the *pedagogic*. The framework was slightly modified by splitting the functional dimension as shown in Fig 6.2. Each of these dimensions has its own fine structure and links back to the type of *assessment task* and the nature of *task assessment*. *Technical* feasibility concerned the extent to which existing technologies were suitable for adaptation to assessment by development of a product or by practical performance and had impacts on equity across different schools and overall costs. *Manageability* concerned the administration of the assessments and the collection of students' work and had impacts on equity and originality of student work. *Functionality* concerned the *validity* and *reliability* of the assessment of students' work created, collected and accessed in digital forms and had impacts on the structure and scoring of the assessment tasks. Finally, the assessment tasks had to be able to support and enhance teaching and learning, that is be aligned to preferred or existing classroom *pedagogy*, if they were to be rightly perceived as feasible under this framework. In this study, students of Applied Information Technology were assessed on the basis of the digital *products* created during an extended digital portfolio and a digital *performance* examination. The focus was on summative, criterion-referenced assessment, capturing the *student's work* digitally and accessing it for *task assessment* to satisfy the *feasibility* and assessment quality guidelines.

What the assessor does in *task assessment* is determined by the *purpose* and *type* of assessment task. *Marking methods* require *marking criteria, rubrics, keys and guides* and *trained assessors* with prerequisite *skills* and *knowledge*. Two methods of marking were employed; *analytical marking* using a rubric and a method of multiple *comparison of pairs* of student work samples. The application of two methods of marking allowed a comparison of the quality of assessment to be achieved. Assessors needed the prerequisite *skills and knowledge* to apply the *marking criteria* to the chosen *marking activities* with sufficient precision to meet the required standards for the course. The *feasibility and assessment quality* indicators address aspects of the *task assessment* process and form the criteria against which success or failure is measured. The *task assessment* process has to be *feasible* in practical and economic terms whilst at the same time being *reliable*, *equitable* and *transparent*. All three aspects must be shown to be equivalent to, or an improvement on, existing methods of assessment.

In the wider sense, *feasibility* was measured with reference to the requirements and aspirations of the key *stake holders*, for ultimately it is the judgement of students, teachers and society at large which will determine the success or failure of this intervention. This assessment system must address all of the caveats and uncertainties discussed above and ultimately be scalable to a large number of candidates if it is to be adopted at state or national level.

Summary

This chapter has examined some of the literature relating to the major themes of Assessment of Human Performance and Computer Assisted Assessment. Beginning with an overview, the chapter progressed to those specific aspects of each domain which had a direct bearing on the study. As a result, a conceptual framework was generated to guide the methodology, data analysis and interpretation for the study. The next chapter will describe the design and method of research, the participants, assessment tasks, data collection and marking.

CHAPTER THREE

METHOD

This chapter describes the research method starting with the wider context of the investigation. The research design and rationale are explained together with details of the participants, the assessment tasks, the data collected and the assessment undertaken. A description of the various forms of data analysis concludes the chapter.

Background

This study formed part of a larger project entitled, *Investigating the feasibility of using digital representations of work for authentic and reliable performance assessment in senior secondary school courses.* The project was a collaboration between the Curriculum Council of Western Australia and the Centre for Schooling and Learning Technologies (CSaLT) of the School of Education at Edith Cowan University, Perth, Western Australia. Four senior secondary courses with large practical/performance components were investigated, with Applied Information Technology being one. The investigation consisted of three phases: *Proof of Concept, Prototype* and *Scalable Product.* It was in the first phase of the larger project that this study was situated.

The broad aim of the present study was to investigate the potential for ICT to support the assessment of practical performance in the external assessment of the senior secondary course of AIT in Western Australian schools. The problem under investigation was how to provide students with the opportunity to demonstrate skills and understanding not amenable to assessment using *pen and paper* but which nonetheless formed a major component of the course. The assessment tasks had to be authentic, able to be scored reliably, manageable and at the same time able to be implemented at reasonable cost.

The study trialled two forms of assessment in AIT; a digital portfolio and a computer-based production /performance task examination, with 115 students studying AIT unit 2B in year 11 and 12 across the seven participating schools (refer to Appendix A for details of the unit). In this first, proof of concept, phase teachers were identified and selected on the basis of their greater experience, computer literacy, technical ability, understanding of the course, involvement in current course development and willingness to participate. Each teacher had at least one AIT class of students either in year 11 or year 12. Student participation in the study was not compulsory but completion of the tasks was because the teacher organised the tasks as part of the course, or at least this was the intention at the outset. Following clearance from the *Human Research Ethics Committee* at Edith Cowan University, students and their parents were required to complete a declaration of informed consent stating their willingness to be involved subject to the usual conditions of ownership and anonymity. Consent was sought to include work output from the tasks to be externally marked and for the students to provide information in the form of surveys and interviews. A panel of trained assessors was chosen, from ECU researchers and moderators from the Curriculum Council, to mark student work output.

At least four visits were made to each school by researchers to introduce the study to students, observe the class at work on the portfolio, to check the technology for the examination and to observe and assist with the examination. During the visits, data and observations were collected and these were written up as soon as possible afterwards into a table of notes and activities. Photographs were taken of the classroom laboratory when it was not in use by the students.

Research Design

The project, of which this study formed a part, is best described as participatory action research. The starting point was the widely perceived mismatch between practice and assessment in AIT shared by teachers, students and assessors and the wish to change things for the better. This study employed an ethnographic, action research methodology using qualitative and quantitative data collected and compiled into multiple case studies (one per school) for evaluation. The findings of this study informed the next cycle of research and centred on an analysis of the perspectives of the participants with data collected for each group.

Rationale for Method

Participatory action research “sets out to explicitly study something in order to change and improve it. It most often arises from an unsatisfactory situation that those most affected wish to alter for the better" (Wadsworth, 1998, p. 6). In this study, the *unsatisfactory situation* was the lack of opportunity afforded to students to adequately demonstrate their skills and knowledge, much of which is of a practical, computer-based nature. Those most affected included the teachers, whose willingness to be involved, to assist in the design of assessment tasks and in the analysis of assessment data, pointed to dissatisfaction with the current arrangements and a wish for improvement. The teachers formed part of a team, which included researchers and representatives of the Curriculum Council of Western Australia, who were collaboratively involved in the design and implementation of the assessment tasks as well as in the collection and interpretation of data arising.

The multi-case approach is described by Burns (1996). With each school making up one of seven separate experiments, any common findings provide compelling evidence for generalisation to a wider population. The approach also allows for refinement and further development of findings based on multiple instances of the same phenomenon under different conditions as described by Willig (2001, p. 82). The study drew on the perspectives of teacher, student and researcher, involved in the same series of activities in similar environments. These perspectives, combined with collection of qualitative and quantitative data, afford opportunity to improve internal reliability and validity of findings by triangulation. The approach drew on the traditions of interpretive research but with the inclusion of some quantitative methods derived from traditional positivist research. This combination is described and advocated by Husen (1994). The blend of the two main paradigms, empirical, quantifiable observation analysed mathematically and holistic, qualitative and interpretative approaches, “are not exclusive, but complementary to each other” (p. 5055).

Target Population and Samples

The study aimed to inform the state-wide implementation of alternative methods of assessment in AIT for which 930 students sat the West Australian Certificate of Education examination in 2008, with the numbers increasing to 1,415 in 2009. The sample under investigation was ultimately selected on the basis of teachers’ willingness to participate and as such is unrepresentative. However the number, size and diversity of the clusters are defensible as a representative sample of the population notwithstanding the bias outlined above. The schools and participant numbers are shown in Table 3.1 with each student being allocated a unique identification code made up of the school code plus a three digit number for example CA106 was the sixth student in class 1 at school CA.

Table 3.1

School type and participant details

School Code	Sector	Student Numbers	Year group
CA	State	20	11
LA	Catholic	10	11
MA	Private	12	11
RA	State	14	12
WA	Private	14	11
XA	Catholic	29*	11
ZA	Catholic	16	11

* 2 classes

AIT was a new course in Western Australia and the teachers in the study had quite diverse interpretations of the course. Typically, teachers were experienced classroom practitioners and had backgrounds in the now discontinued, wholly-school-assessed courses of *Business Information Technology* and *Digital Media*. Five were male and two female. A typical student was a year 11 male interested in computers from a practical rather than a theoretical standpoint and broadly classifiable as non-academic. Comfortable around computers and computer confident, these students sometimes studied other computer-based courses in addition to AIT. In a typical lesson, students frequently switched between time on task and other activities such as completion of other schoolwork, web surfing, games, e-mail and music where network restrictions permitted. Selection of AIT was not common amongst those with high aspirations of tertiary entrance. It was often used as a subject to make up a complement and gained numbers from students dropping out of other courses when these proved to be unexpectedly difficult.

Expressions of interest were sought from teachers offering the course and seven were selected to participate. Of the seven, two were state, three Catholic and two from the independent sector. Each teacher had at least one class of senior secondary students. Six of the classes comprised entirely year 11 students and one entirely year 12 students. The students were wide-ranging in their ability and socio-economic status. Teachers were all experienced practitioners with a personal interest in the development of innovative assessment methods in AIT.

Context of Study

The context of the study was the Western Australian course in Applied Information Technology unit 2B. The course was introduced in 2007 with the syllabus last updated in June 2009. In the introduction to the syllabus for unit 2B, the following statement is made:

> The focus for this unit is information and communication technologies in business. Skills, principles and practices associated with various types of businesses to enhance students' career prospects are emphasised. Students examine the use of ICT in a range of administrative and business environments. They identify and explain the components and configuration of a computer system to meet the needs of the organisation. Students design information solutions for problems encountered in these contexts and understand the social issues inherent in work practices. (Curriculum Council of Western Australia, 2009, p. 6)

Specific reference is made to word processing, presentation software in business, simple spreadsheets, business office suites, publishing and creative application of information design principles and elements in the creation of business related technology products.

Table 3.2 shows the types of assessment and the range of acceptable weightings stated in the syllabus. The course is designed to be completed in 55 class contact hours. Assessment is school based, unless students are in their final year of schooling when those students who are studying at least one Stage 2 pair of units (e.g. 2A/2B) must sit a three hour written examination in this course, unless they are exempt. The "examination will assess the specific content, knowledge and skills described in the syllabus for the pair of units studied" (Curriculum Council of Western Australia, 2009, p. 9). School based assessment is of three types: Investigation, Production/performance and Response as shown in Table 3.2 and centred on four outcomes:

i) technology process,
ii) understanding information and communication technologies,
iii) quality of information solutions and

iv) information and communication technologies in society.

In this study, teachers were requested to include, as part of their assessment for the course, the two activities developed for the purposes of research, namely an extended portfolio task and a computer-based examination.

Table 3.2

Assessment Type and Weightings for AIT Stage 2

Weighting	Type of assessment
20–30%	Investigation
	Research works in which students plan, conduct and communicate an investigation. Investigation of ICT-related issues or cultural contexts, exploring a range of primary and secondary sources. Best suited to the collection of evidence of student achievement of Outcomes 1, 2 and 4.
50–60%	Production /performance
	Extended production project in which students explore ideas and control the processes required to manage the quality of production. Students engage in an activity or on-the-spot evaluation of a performance. This may be one large production /performance task or it may be two or more smaller tasks. Manage a range of production processes, evaluating and modifying them as necessary. Demonstrate an understanding of styles, structures, codes and conventions and the development of confidence and competence in the use of technologies, skills and processes in a range of contexts. Types of evidence may include: a journal to show evidence of exploration and the development of ideas, reflection on learning processes and critical evaluation and modification of ideas, portfolios and products. Best suited to the collection of evidence of student achievement of Outcomes 1 and 3.
20–30%	Response
	Students apply their knowledge and skills in analysing and responding to a series of stimuli or prompts. Response to, analysis and evaluation of own or professional information technology products. Types of evidence may include: observation checklists, journal and evaluation tools (self or peer). Best suited to the collection of evidence of student achievement of Outcomes 2 and 4

Role of the Researcher

The researcher was part of the task development team, observed the classes in action, interviewed the students, invigilated the examination and ran the survey. The researcher was one of the two independent analytical markers and one of the five independent *comparative-pairs* markers. One of the case studies (XA) was from the researcher's school but was taught by another teacher. This independent but involved role assisted with an understanding of the quality of student work, the problems encountered and what the students had done.

Assessment Tasks

This section describes the assessment tasks developed for the study and the processes used. These comprised two main forms: a digital reflective process portfolio and a computer-based examination.

Development of Tasks

It was important that the assessment tasks constituted good professional practice, met the requirements of the course and were reasonably simple to implement in a real school by a ‘good’ teacher. The starting point for the development of the tasks was the *Situation Analysis* set out in Appendix B. The analysis looked at what might reasonably be possible within the requirements of the course and the constraints of the school environment paying due regard to the capabilities of the students and teachers. The *Rationale for Assessment* (Appendix C) was based on the situation analysis, the AIT syllabus and informed by the guidelines for digital assessment set out in the review of literature in Chapter Two. This combination set out the scope and structure of possible assessment tasks and formed the starting point for authoring the tasks themselves.

Portfolio

The portfolio was developed by the team of teachers, researchers and Curriculum Council officers with reference to the *Rationale for Assessment*, course syllabus for AIT unit 2B together with the principles derived from a review of literature. The portfolio comprised three components completed over an extended time period with all work recorded digitally. In component one, students were required to design and create an information technology solution to meet the requirements of a design brief. An *example*, with full documentation ready for implementation, was supplied, the design brief consisting of the establishment of a web presence for a teenage clothing retailer. The second component required students to produce a design process document detailing the investigation, design, production and evaluation of their solution to component one. The third component of the portfolio required the student to select two further digital artefacts, different from each other and from the portfolio prototype (component one) which had been created earlier in their course. The intention was to allow students to demonstrate the range and depth of the relevant practical skills they had acquired. Each artefact was to be accompanied by a brief description of the software used and skills employed in its creation. The complete description of the portfolio assessment task is included in Appendix F.

Computer-based Examination

The examination was developed by the team of teachers, researchers and Curriculum Council officers with reference to the *Rationale for Assessment* (refer to Appendix C), course syllabus for AIT unit 2B (Appendix A) together with the principles derived from a review of literature. The task was made up of two components. The first was a one-hour theory section, entirely independent of the second. Students were asked to respond to a series of reflective questions by word processing their answers into a document template provided on a USB flash drive. The questions were designed to draw out details of the student's understanding of the technology process in relation to the product development undertaken as component one of the portfolio above.

The second part of the examination involved the design and creation of a business logo and a tri-fold advertising brochure for a resort hotel. Unformatted numerical data about the hotel was supplied and some of this had to be graphed, tabulated and ordered. Students were required to add the logo, manipulate some of the 18 image files supplied and write the text to promote the hotel and caption the photographs to complete the product. Finally, students were asked to plan and record an audio reflection of their work explaining a little about how it was done, pointing out, with justification, the design principles and conventions used. The complete description of the examination assessment task is included in Appendix G.

Implementation of Assessment Tasks

Teachers were asked to implement both the portfolio and the examination with their classes as part of their regular programme.

Portfolio

It was intended that the portfolio be implemented fully in class time with students having 15 hours over four weeks to develop a prototype by applying the technology process to the design brief. Bearing in mind the voluntary nature of participation by teachers, it was not possible to enforce compliance with the tasks developed for the study even though this was the expectation at the outset. Indeed, flexibility in the choice of context for component one of the portfolio was the intention with each participating teacher being allowed to develop their own design brief as required. There were some differences in the contexts chosen by each of the seven schools involved, though four of the seven used the example task supplied, namely an e-commerce website for a teenage clothing retailer. The other three used a budget airline booking website, an Olympic Games bid website and a training animation. The intention was that the work should be entirely that of the student and done in class time, although whether or not this was the case could not be ascertained accurately.

Students were requested to record the stages of development and this information was drawn upon in the second component of the portfolio, the creation in class over five hours of a design process document. The document was based around a template supplied and was designed to lead the student through the technology process of investigation, design, production and evaluation in relation to the portfolio product.

Component three, the two digital artefacts and their supporting documentation were left entirely to the discretion of the teacher and student. Teachers collated and submitted all work on disk for all students for whom completed and signed consent forms had been received.

Computer-based Examination

Execution of the examination was relatively consistent across the seven cases with the researcher assisting the teacher in managing the implementation over a three hour period. Students were provided with an examination paper, two double sided A4 design sheets, a headset and microphone and a four gigabyte USB flash drive containing a data file and 18 preselected digital photographs appropriate to the task. Each USB device was labelled with the student's identification code. Soft copies of the design sheets were also included to give students the option to design on computer if this was their preference. In six of the seven schools, the examination ran continuously for three hours with 10 minutes reading time and in one, the two parts were completed separately with only the two hour practical section being observed.

In the initial configuration of the examination, the one hour theory section (keyboarded answers to questions) was placed at the beginning. This presented an unforeseen problem which became evident at the first implementation when students finished early and were not allowed to go onto the practical section. As the two sections were independent, it was subsequently decided to reverse the order, placing the two hour practical section first and this seemed better suited to the time allowed as even where students had finished, there was a willingness to re-work and refine their brochure and so invigilating problems which followed from students sitting idle, were eliminated. The reversal of the parts now meant the audio reflection, the concluding part of the practical, came in the middle of the examination and this was much harder to supervise as students were in effect allowed to talk during the examination.

A typical implementation began with an audio test. Students were asked to open a suitable application and test the capability of their computer in recording a short audio clip. Reading time followed during which students were allowed to browse the files on the USB flash drive. The intention was that students would check and alert the invigilator to any problems or omissions. All parts of the production process were captured digitally and saved to the USB flash drive with the exception of the design, where students had the option to develop their ideas using application software or to develop ideas on paper and have these collected at the end for scanning.

In completing the task, students were allowed to use any software available to them, but were restricted to the data and image files supplied and could not add to these by downloading or retrieving from personal storage. At some schools, internet access was disabled but at others students could have browsed for ideas or even downloaded though students were told this was not allowed. The time allocated to each part was strictly imposed and students were not permitted to go back to the practical if they completed the written section inside the hour allocated. Students who finished early were given a questionnaire to complete.

Data Collection

This section describes each of the types and sources of data and the respective method of collection. Table 3.3 shows the subsidiary research questions and the data sources used to address them.

Classroom Observation

At each visit, field notes were compiled. These included records of conversations with teachers and students, details of the setting (photographs were taken of each venue), observations and impressions. Reflections on the visit were added as soon as possible after the visit; typically in the evening of the day of the visit.

Table 3.3

Data Sources Related to Subsidiary Research Questions

Research Question		Data Sources
1. What are the advantages of digital capture of students' performance in support of summative assessment of practical ability in the senior secondary AIT course?		Observations from school visits, student surveys and interviews, teacher interviews and assessor interviews
2. What are the limitations of digital capture of students' performance in support of summative assessment of practical ability in the senior secondary AIT course?		Observations from school visits, student surveys and interviews, teacher interviews and assessor interviews
3. How feasible is the digital capture of students' performance in different forms of summative assessment in AIT with respect to:	Manageability	Researcher observation, student questionnaires, student interviews and teacher interviews
	Technical facility	Observations from school visits, student interviews, teacher interviews and assessor interviews
	Functionality	Interviews with teachers and assessors, assessors' scores awarded
	Pedagogy	Observations from school visits, questionnaires and interviews with students and teachers
4. Do judgements by multiple comparisons of pairs, produce reliable scores when applied to summative assessment of practical performances in the senior secondary AIT course?		Assessors' scores from marking
5. Which method of marking, analytical or comparative pairs, was better in assessing student practical performance in AIT?		Assessors' scores from marking. Assessor interviews. Teachers student performance data

Student Survey

All students completed a questionnaire consisting of 70 closed response items and two open response items. The questionnaire is set out in the Appendix H. The questionnaire sought students opinions on the examination itself (items E1(a) to E2(k)) and the the portfolio of work completed during the term (items P1(a) to P2(k)), use of computers and other digital devices (items 5 to 10(e)), attitudes to using computers (items 11(a) to 12(f)) and facility with computer applications (items 13(a) to 13(k)). In all cases it was possible to collect the completed questionnaire from students immediately after the examination. Typically, many students finished the second part of the examination, the reflective questions, well inside the one-hour time allowed and were able to complete the questionnaire whilst waiting for the examination to conclude. For each school, the summary of responses to the student questionnaires was collated into a spreadsheet. Numerical values were assigned to the closed responses according to a predetermined code. These spreadsheets were imported into SPSS (Statistical Package for the Social Sciences) for statistical analysis.

Student Forums

At all schools, a student focus group of between four and six students was convened immediately after the conclusion of the examination with the intention of ascertaining their views on the nature and complexity of the assessment task. The forums were semi-structured using the set of open-ended questions which are set out in the Appendix I. The discussion aimed to address the dimensions of manageability, functionality, pedagogy and feasibility using semi-structured interviewing. Students were encouraged to comment on any aspect of the examination and to suggest areas of improvement. The discussions were audio recorded and later summarised with key points being transcribed.

Teacher Interviews

Teacher comments and suggestions were noted during the observation visits. Teachers were asked to share their views and experiences pertaining to the nature, organisation and delivery of the tasks and these formed part of the field notes for each case. At the conclusion of the study, teachers were requested to complete an *e-mail interview* the form of which is set out in the Appendix J. This was issued to and completed by all participating teachers. The results of each teacher interview as well as the notes obtained at each visit were summarised and added to each case study.

Student Work

Collection of student work from the examination simply involved collecting in the labelled USB flash drive and up to four labelled design sheets. Collection of student work for the portfolio was more problematic. Since all teachers were volunteers and their goodwill vital to the success of the study, it was not considered appropriate to mandate a collection method or deadline. This introduced several difficulties with, for various reasons, many student portfolios being incomplete or missing. Where student portfolios were received, they usually came on a single DVD for each school with a folder for each student. File types were checked for compatibility with the marking system, modified where necessary and uploaded to the web based repository for marking.

Achievement Data (Teacher Marks)

Teacher marks for students for the *semester* (ideally including the portfolio) and for the examination were requested. There were again problems in acquiring these and even when received, deducing to what exactly they referred. The intention was that the portfolio task would form part of the assessment structure for the course and be added to other results to generate a final semester mark. Similarly, it was intended to have the examination marked by the teacher, using criteria or method of their own choice, and for this to be submitted. What was received varied from school to school and ranged from complete compliance with intention to total absence of marks.

External Marking

Two methods of external marking were employed; analytical numerical marking guided by rubrics and comparative-pairs marking guided by criteria statements. Analytical marking for all participating students was completed independently by two computing teacher experts acting as assessors. Files for the portfolio (product, process document and artefacts) and examination (theory and practical) were uploaded to a database from where they could be viewed through a web browser from any computer with an internet connection. Access to the files was password protected. The opening screen showed students listed in order by school with various browse and select options. Selection of a student code opened a split screen with the first marking rubric (portfolio product) to the left and a window to the right in which linked files making up the product could be displayed. Most parts of the portfolio opened directly in the right panel; those which did not had to be downloaded and opened locally. The marking was done by selecting one of a number of radio buttons corresponding to marks for each aspect of the work and totalling these was automatic. Tabs for each of the five pieces of work allowed progression to the next item with display of the associated marking rubric and hyperlinks to the content. Screen shots of the marking tool are shown in Figures 4.3 and 4.4.

Comparative-pairs marking was undertaken by five computing teacher experts (two having been involved in the analytical marking) for the practical component of the examination only. A reduced sample of 60 students was selected by eliminating students whose examination content was incomplete. Students who did not have an audio reflection or for whom the scanned design documents were incomplete or difficult to read were removed. Further reduction was achieved by removing 12 students whose scores in the analytical marking were similar to many other students. In this way, a good range of the exemplars was assembled each of which contained a complete set of the components of the examination. One holistic and three specific criteria were developed from the task specifications and these guided the markers through a series of predetermined paired comparisons. The marking tool comprised a web based database holding links to the work samples allowing pairs of student work to be displayed side by side with the four assessment criteria displayed between them. Checkboxes recorded the assessor’s preference for one piece of work over the other according to each of the four criteria. A button advanced the system to view the next pair with the system storing the result of each comparison. A text field allowed markers to record comments on a particular work sample and these were stored with the sample and reappeared when that sample was again involved in a comparison. Screen shots of the marking tool are shown in Figures 4.4 and 4.6.

Data Analysis

The collected data for each case were first analysed separately. The results of these were then combined for all cases and further analysed using the Feasibility Framework for Applied Information Technology with a multi-case study approach. Some data were analysed for the whole sample, for example student surveys and analytical marking scores, to address the research questions from the perspectives of the four dimensions:

i) Manageability. Researcher observation, student questionnaires, student interviews and teacher interviews were analysed to determine if the assessment task was physically possible in a standard AIT classroom in a senior secondary school with a typical student cohort.
ii) Technical**.** Observations from school visits, student interviews, teacher

interviews and assessor interviews were analysed to expose any technical limitations to the adaptation of existing technologies to digital data capture, collation and online assessment.

iii) Functional. Interviews with teachers and assessors, assessors' scores awarded and the examination of student assessment data on other tasks in comparison to the work on the assessment task, were analysed to provide evidence of validity, reliability and comparability.

iv) Pedagogic. Observations from school visits, questionnaires and interviews with students and teachers were analysed to provide evidence in support of the use of the task from an educational perspective. That is, the task should be more than merely an assessment tool and should also have the potential to reinforce, enhance and deepen student learning in AIT.

The Feasibility Framework included two additional factors: constraints and benefits. The specific analytic method depended on the type of data.

Classroom Observation

Field notes were read, re-read and categorised according to the headings in the feasibility framework. These coded observations, comments and interpretations were summarised in the case study for each school.

Student Survey

Questionnaire data were collated into a spreadsheet. The responses were analysed to produce frequency and descriptive statistics for each school and the population as a whole using the computer software SPSS. Responses were coded numerically to generate seven scales which were derived from combining selected items from the questionnaire. The scales were derived from the questionnaire and were tested beforehand in a pilot study for the main project. The definitions of the scales are given in Table 3.4. The *eAssess* and *eAssessP* scales were the perceived suitability and perceived efficacy of computer use for the examination and the portfolio. The *Apply* and *Attitude* scales were measures the application of computers to everyday tasks and sentiment towards computers. The *Confid* and *Skills* scales were measures of students' confidence using computers and self-assessed skills with everyday applications. Since the scales for *eAssess, eAssessP* and *Skills* consisted of four discrete values (1, 2, 3 and 4) the midpoints are at 2.5. For *Apply, Attitude* and *Confidence* the sales range from 1 to 3 with mid-points of 2.0. A seventh scale *SCUse* was a student's estimate of the average time spent using a computer at school in minutes each day

Table 3.4

Definition of scales derived from the student questionnaire

Name	Description
eAssess	Perceived efficacy of computer use for the examination (combination of all items in Question E2). Potential range of responses between 1 and 4. Mid-point 2.5.
eAssessP	Perceived efficacy of computer use for the portfolio (combination of all items in Question P2). Potential range of responses between 1 and 4. Mid-point 2.5.
Apply	Application of computer to various uses (combination of all items in Question 10). Score between 1 and 3. Mid-point 2.0
Attitude	Attitude towards using computers (combination of all items in Question 11). Potential range of responses between 1 and 3. Mid-point 2.0
Confidence	Confidence in using computers (combination of all items in Question 12). Potential range of responses between 1 and 3. Mid-point 2.0
Skills	Self-assessment of ICT skills (combination of all items in Question 13). Potential range of responses between 1 and 4. Mid-point 2.5.
SCUse	Estimate of time in minutes per day spent using computers at school (combination of all items in Question 8).

Students were asked to list the two best things about doing the portfolio and examination by computer and to list the two worst things about doing the portfolio and examination by computer. Responses to these open-response items were typed into a spreadsheet for each school and any repeating patterns in responses were tallied to assist in drawing out themes. The more common responses were summarised in each case study.

Student Forums

Audio recordings of each interview session were played back in short sections and summarised for each case. Any points of consensus, incongruity, anomaly or innovation were transcribed.

Teacher Interviews

Interview data were summarised for each case to assist in the development of themes. Points of consensus, criticism and suggestions for improvement were summarised for each case and for all cases. These data were confirmed with teachers through provision of case study reports to each school.

Student Work and Scores from Marking

The two assessors produced numerical scores for each of the three portfolio components and two examination components for all students. Descriptive statistics for these scores were calculated for all students. The scores awarded were tested for reliability by calculating correlation statistics on the total mark awarded as well as the marks awarded to individual sections. Where data existed, correlation coefficients between individual assessors and teacher marks and assessor average and teacher marks were also calculated. Teacher marks comprised the final semester mark and the mark awarded for the examination. Students were also ranked according to the total score awarded by each external assessor and these were also correlated. The rank was the position out of the whole sample of 115 students spread over seven schools. For example a *rank assessor average (Ass Ave)* of 86.5 meant that this students score ranked 86th equal out of 115. The *rank teacher's semester (Tch Sem)* was the position of the score awarded by the teacher compared to the scores awarded by all other teachers in the sample. For example, a teacher rank of 18.0 meant that the score awarded by the teacher was the 18th highest out of 115. Correlation statistics were also calculated with respect to students' ranking as determined by marks awarded by the teacher and the marks of external assessors both as individuals and on average. Further analysis of the results of the analytical marking was completed by applying a Rasch polytomous model using the *RUMM 2020* software (RummLab, 2011).

The comparative-pairs marking process considered only the practical component of the examination for a sub-set of 60 students selected on the basis of having complete data sets. A Rasch dichotomous model using the *RUMM CC* software produced an interval scale allowing a mark to be assigned based on the relative merit of the work. This mark was correlated with the mark obtained by marking the same task by the analytical method for each assessor individually and for the average of the two. Further comparisons between analytical and comparative-pairs marking were made for marks and rankings obtained from teacher generated data where available. The Rasch model provided a basis and justification for placing a student in a particular location on a continuum according to the total score awarded in the analytical marking.

Scores derived from the analytical marking by the assessors were imported into SPSS from where summary and descriptive statistics were generated. This allowed for an analysis of differences between the classes to be made. Correlation coefficients were generated to allow comparison to be made between the scores awarded by the assessors for each of the following situations:

i) Total scores awarded by each assessor for *all* components for *all* students
ii) Total scores awarded by each assessor for *all* components for *each class*
iii) Scores awarded by each assessor for *all* students for *each* component.

In addition, a score was generated from the ranking obtained from assessment by the method of comparative-pairs by five assessors for component 5, the practical section of the examination, for 60 selected students' work. These scores were analysed for their inter-rater reliability by calculation of a separation index. Finally, Rasch analysis of the portfolio and examination was undertaken using the scores produced by the analytical marking of two assessors.

Assessor Interviews

Feedback was obtained from informal discussion between the researcher and two of the assessors (one analytical, one comparative pairs) *after* the conclusion of the marking process. The researcher, being involved in both types of marking, was a major source of data pertaining to the assessment process.

Summary

This chapter described the research methods, data sources, collection, organisation and storage of data. The nature of each data source was described as was the approach to data analysis and interpretation. The following chapter looks at the analysis of the data collected as a whole across all participating schools.

CHAPTER FOUR

RESULTS OF DATA ANALYSIS

Introduction

The results of the analysis of data related to the implementation of the AIT assessment task are presented in this chapter, starting with a discussion of the implementation of the task and the technologies employed. This is followed by the results of analysis of the data collected from surveys and interviews of students, teachers and assessors and from marking of the students' work. Results for each school are not discussed separately as these are presented as case studies in the subsequent chapter. This chapter provides the context for the case studies by giving an overview of the full dataset.

The AIT assessment tasks were implemented at seven schools with a total of eight classes (one at each school, two at XA) of Year 11 or 12 students studying the AIT Unit 2B. This involved a total of 115 students, seven teachers and five assessors (two for the analytical marking and an additional three for the pair-wise comparisons). For each of the eight classes the three components of the portfolio and two components of the computer-based examination, described in Chapter Three, Appendix F and Appendix G, were incorporated to some extent within their second semester programme.

Task Implementation

This section provides background analysis across the classes on the implementation of the assessment task, drawing on researcher observation and the interview data. Although there were some differences in the manner in which the assessment tasks were implemented for each of the eight classes of students, in most ways they were similar. Each class was visited at least four times during the completion of the five components of the assessment task. All of the sessions were conducted by the teacher in a computer laboratory at the school. For the examination session, the researcher assisted each teacher in invigilation. Table 4.1 shows the number of submissions by students of portfolio and examination work and the percentage of portfolios and examinations submitted for each class. There were a variety of reasons behind any omission of students' work from the study. For example in one school (CA), parents of three students withheld permission to take part in the study. In another school (WA), despite repeated requests to the teacher, 14 portfolios, although completed, were not delivered to the researcher. One student (LA) was suspended from computer use for disciplinary reasons and could not complete the examination. In spite of these omissions, 96% of students completed the examination and 84% of students had a portfolio submitted.

Table 4.1

Number of students involved by school and number and percentage of submissions of tasks.

		Portfolio submissions		Examination submissions	
School	Students	Number	%	Number	%
CA	23	20	87	20	87
LA	11	10	91	10	91
MA	12	12	100	12	100
RA	14	14	100	14	100
WA	14	0	0	14	100
XA	29	29	100	29	100
ZA	17	16	94	16	94
Total sample	120	101	84	115	96

Portfolio Product Development

The first component of the portfolio (component 1) was the prototype product developed in response to a design brief. Each teacher was permitted to set their own design brief for the portfolio product although four (CA, LA, RA, ZA) used the example, The *Miss Shoppe* website, provided with the project documentation reproduced in Appendix F. Students were set 15 hours of class time over 4 weeks to develop a prototype. It was intended that all work should be completed in class but teachers varied in the extent to which this was enforced. Hardware and software were restricted to those available at the school.

The focus of the activity was the application of the whole technology process to a real-world context, as set out in the scenario contained in the design brief. Students were informed by the project documentation that the product was required to:

- suit the intended purpose and audience/users
- meet the requirements of the design brief and/or client specifications
- illustrate creative application of information design principles and technologies

- make use of appropriate information structures, forms, layouts and symbols
- employ relevant standards and conventions to create multi-tiered information solutions
- use appropriate methods and techniques to represent the design of information solutions

Portfolio Design Process Document

The second component of the portfolio (component 2) was the design process document which students developed from a template supplied with the task documentation. On completion of the product, evidence of the investigation, design, production and evaluation processes undertaken was to be collated into a *Design Process Document*, this forming the second component of the portfolio. This was in four sections (Investigation, Design, Production, and Evaluation) with prompting questions and page limits for each. Students were to be given 5 hours of class time to collate the document using material created during their product development. At least two teachers (CA, XA) permitted students to complete this at home and during visits some collaboration was observed although this was not the intention. There was widespread evidence of this component being produced in the manner described by Ridgeway et al. (2004, p. 28) in which "students create some artefact, then 'back-fill' by inventing the development process post hoc" rather than drawing directly on existing work. The researcher observed this occurring and recognised the results in some submissions of the portfolio.

Portfolio Extra Artefacts

The third component of the portfolio (component 3) invited students to showcase any additional skills acquired during the course by the submission of two further digital artefacts. A template, reproduced in Appendix F, was supplied with the project documentation to lead students through a description of the hardware and software used, the techniques and skills involved and any help received from others. It was intended at the outset that students would select the artefacts from work done earlier in the course and all teachers indicated that students would be able to do this.

Examination Part A: Reflective Response Questions

The first part of the examination (component 4) consisted of a series of structured reflective questions which sought to elicit students' recall and understanding of the technology process as applied to the creation of the portfolio product, portfolio component one. Student responses were typed into one of thirteen text boxes provided below each question in a *MS Word* document saved to the USB flash drive which was issued to each student before the examination. This first part of the examination was allocated one hour and typically was done with the second part, but was not required to be so. In one of the seven cases (ZA), the reflective response questions were done on a different day; in another (RA), the first implementation of the examination, the reflective response questions were done first. In the remaining five, the reflective response questions were completed immediately after the examination part B. The decision to swap the order was based on the experience gained at the first implementation (RA). Students finished the reflective questions well before the time allowed and had to wait to commence the next section. The waiting wasted time and introduced invigilation difficulties and it was decided to reverse the sections in all subsequent implementations. With the sections reversed, any student who finished early could complete the questionnaire and leave the examination. This minimised disruption and used the time more efficiently.

Examination Part B: Production

The second part of the examination (component 5) consisted of a production /performance task in which students were given a limited, real-world design brief and prompted to follow the technology process to create a digital product. The design brief, set out in Appendix G, called for the creation of a marketing brochure for a resort in a delicate conservation area. Two hours were allocated to this task and typically this was done as the first part of the examination, but being entirely independent, was not required to be so.

With the exception of design sketches, which had the option of being paper or computer-based, the entire examination was done on computer, students' responses being saved as digital files in various formats. Students were given a paper copy of the examination, a 4GB USB flash drive and an audio headset with microphone to record an audio reflection. There was 10 minutes reading time prior to the commencement of the three- hour paper which was completed under examination conditions with the teacher and researcher invigilating. The USB drive contained 18 digital photographs, a text file of data, design templates in *MS Word* and *PowerPoint* and a template for preparation of an audio reflection in *MS Word*. A hard copy of the design template in *MS Word* was also supplied to give students the option of designing on paper. Students were permitted to use any software available on their desktop computer saving their work to the USB drive. Students were not permitted to continue with the practical once the two hours allocated had expired.

Technologies used by Students

All student work, apart from the initial design section of *Examination Part B*, had to be submitted in an appropriate digital form using computer systems. The design sheets were scanned into PDF files by a research assistant. For the production component of the portfolio students might have also used other peripheral devices such as cameras depending on the nature of the design brief. The implementation of the examination components of the assessment task required the use of a computer workstation, a USB flash drive, a headset with microphone and appropriate software (office and graphics applications). The teacher at the school was responsible for setting up the workstations while the researcher provided the USB flash drive and headsets.

Collection of Student Work and Creation of On-Line Repository

For the computer-based examination all student digital work was saved by the student to the USB flash drive allocated to the student and typically a copy was also saved to the school's server. Students' design work, that was done on paper, was collected and either scanned or photographed to add to the digital work in preparation for marking. Typically, student portfolio work was provided by the teacher on a disk and organised by student folder. All digital work was transferred by a research assistant to a folder, named using the student's ID, on a server at Edith Cowan University, so that the work could be accessed by assessors. Each student folder contained sub-folders for the portfolio product, the artefacts and the examination as shown in Figure 4.1.

The portfolio process document was a PDF file placed within the main folder on its own. Within each sub-folder there was an index page that was used by the online marking tools to display the contents of the folder and this page contained links to the other files. The artefact folder contained a PDF file (ArtefactDescription.pdf) of the student's descriptions of the two artefacts as well as the artefacts themselves. The examination folder contained all the files copied from the USB flash drive used in the

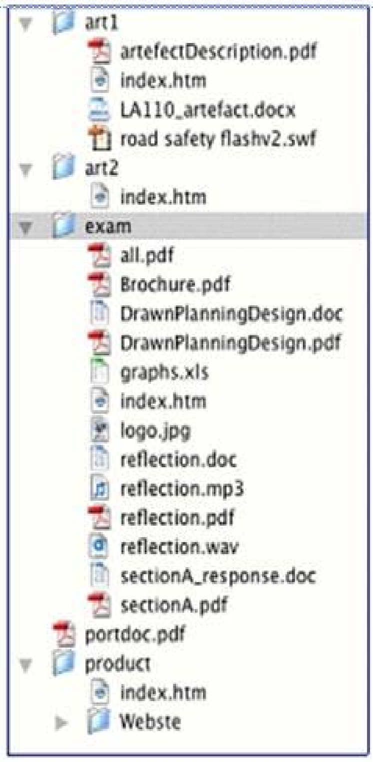


Figure 4.1 On-line files repository for a single student

examination and PDF versions of the brochure, reflections, plans, examination section A responses and one PDF file with all the examination section B files combined.

Summaries of Results of Analysis of Data

This section describes and presents the results of data analysis for each source of data for the whole sample. In the following chapter the results are presented for each class as case studies. Data collected for analysis comprised the following:

- observation of the classes in action; informal discussion with the students and teachers;
- survey of students by questionnaire;
- interviews with teachers and a small group of students;
- interviews with assessors;
- results of analytical, rubric–based marking; results of marking by comparative pairs; and
- results of marking by class teacher and overall semester mark.

The results of analysis of each source of data are now presented separately for the whole sample.

Observation of Classes

Participating teachers had responded to an expression of interest and had been provided with full documentation of the intention of the study and a complete set of student support materials. At the time of the first visit, students were already at work on the first portfolio component, the prototype product. Observations at this stage were of the class, not of individual students, so no record was kept to identify specific participants.

Observation of Portfolio Product

Typically, the teacher introduced the researcher with an invitation to explain the purpose and structure of the study and to field any questions. Students were always attentive and appeared keen to be involved. Much of the remainder of the visit was taken up with ‘over the shoulder’ observation of the students at work, often leading to discussion of details of techniques and software used. Students were usually eager to show and tell, particularly when a feature of their work displayed originality and sophistication. Many went to great lengths to perfect a ‘cool’ feature (such as a roll over button) and it was obvious that completing the whole product was not always a high priority. There was little evidence of students working to a plan, of formal time management or journals. Generally, students developed their ideas on the computer, selecting colours, shapes and effects by trial and error. During, or immediately after the class, the researcher and teacher discussed potential problems and this often led to a wider discussion of the study, the course itself and the philosophy of assessment in AIT. Photographs of the computer room were taken and plans were made to resolve any technical problems in preparation for the examination. Students had the opportunity to share and discuss their work and although no direct observation of collusion was observed, it could not be ruled out. Further, students had the opportunity to take their work outside the school environment and outside the control of the teacher.

Observation of Portfolio Process Document

Once again, it was usual for the teacher to introduce the researcher and in two schools (RA, CA) the researcher was asked by the teacher to go over the requirements of this part of the portfolio with the requirements of this section being displayed using a data projector. This component was essentially a written task, collating and documenting the technology process for the product and should have been straightforward if investigation, design and production documentation had been created and saved previously, leaving only the evaluation to be done. Generally, it appeared that this section was often completed retrospectively. For example, production plans and timelines, which should have been produced prior to commencement of product development, were now needed, and these had to be ‘made up’. Documentation relating to investigation was also typically sparse. Students often said they had visited several websites but failed to document any analysis of strengths and weaknesses. Some of the questions were misunderstood or misinterpreted and there was a good deal of repetition. For example, evaluation *criteria* were confused with evaluation *method*, and alternative solutions (that is the student’s own designs) were confused with existing solutions, usually related websites a student had visited. There was an impression among students in many schools that this was not what AIT was supposed to be. One student (LA) said, “I know what a website should look like…no need to design it”.

Observation of Examination

The three hour examination consisted of two distinct parts; a one hour theory section and a two hour practical section. In six of the seven schools, these were completed concurrently, the practical section preceding the theory section, in five of these. In one school the theory section was completed separately and only the practical section was observed. At the first implementation (RA), the theory section preceded the practical but this led to problems. Many students finished the one-hour section after 20 minutes and sat waiting to commence the practical section. During this time two students became engaged in a computer game and another in solitaire. There was a lot of looking around and invigilation was difficult. For this reason, subsequent implementations placed the theory section second so that if completed early, students could use the time for the questionnaire or be dismissed from the examination room. This made invigilation much easier as students were fully engaged for the two hours of the practical. The only disadvantage was the recording of the audio reflection during the last five minutes of the practical section. Students were necessarily permitted to speak during this time and invigilation was made difficult by the amusement borne of the self-conscious excitement which students felt during this novel experience. This coupled with widespread hardware malfunction led to less than satisfactory compliance with traditional examination standards.

One major shortcoming was the visibility of students work. Although students were seated as far apart as possible, their screens could easily be seen by neighbours and it was observed that ideas sometimes appeared to diffuse laterally though this was relatively rare and might have been coincidental. In general students were fully engaged with the practical section and many times the question of returning to the practical after completing the theory section was raised. This was always declined but did present a further invigilation problem.

Survey of Students

Across all classes, 110 students (81 male, 29 female, 96 in year 11 and 14 in year 12) completed a questionnaire consisting of 70 closed response items and two open-response items. Responses to the open-response items were tabulated to assist in drawing out themes and a summary of the main points is shown in Table 4.2.

Table 4.2

Summary of the Best and Worst Aspects of Computer-based Portfolio and Examination

The best things …	The worst things …
Portfolio: Students overwhelmingly commented on the ease and enjoyment of working on the computer. The exact aspect of easiness cited varied, but frequent mention was made of editing, correcting errors, speed of action and physical comfort. Some students mentioned the creative freedom afforded to them and others specific parts such as filming or learning new skills.	Portfolio: By far the most prevalent comments referred to the fear, rational or otherwise, that it could all go horribly wrong. These fears were often mentioned together with a critical appraisal of the hardware provided, and clearly for many schools, the software employed by students and the manner in which it was deployed strained system resources. The time taken to become familiar with new software and processes was also mentioned.
Examination: Students again commented on the ease and speed of working on the computer. The exact aspect of easiness cited varied, but frequent mention was made of speed of working, not having to write, error correction, speed of writing, amount of writing, speed of action and physical comfort. Several comments alluded to the reduced stress levels experienced in this form of examination	Examination: Students again cited the potential of losing work, made frequent mention of the perceived inadequacy of hardware and software particularly those related to speed such as lagging and not-responding. There were also frequent comments about distraction caused by noise of other students typing and time allowed; paucity for completion of the practical task, surfeit for the theory section.

Broadly, the questionnaire sought students' opinions on the examination itself, the portfolio of work completed during the term, the use of computers and other digital devices, attitudes to using computers and respondents' facility with computer applications. A copy of the survey may be found in the Appendix H and responses by case in Appendix K. The minimum, maximum, mean, standard deviation and frequency of responses were calculated for each closed response item using SPSS. A number of scales were derived by combining responses to items from the questionnaire. These scales were derived from translation of the *Likert* scales into numbers. For example, in question E2(a) "It was easy to use the computer for doing the exam" the responses *Strongly Agree*, *Agree*, *Disagree and Strongly Disagree* were coded four, three, two and one respectively. The numerical scores were aggregated with those of similar question types to produce scores on scales related to six constructs. These were, ease of completion of an examination on a computer (*eAssess*), ease of completion of a portfolio by computer (*eAssessP*), frequency with which computers were used to perform common tasks (*Apply*), attitude to using computers (*Attitude*), confidence in using computers (*Confidence*) and self-assessment of computer skills (*Skills*). Definition of the scales was given in Table 3.3 in Chapter Three. A seventh scale, student computer use, comprised the aggregation of students' estimate of the amount of time in minutes spent working at a computer at school each day in a typical week. Some descriptive statistics for these scales are shown in Table 4.3 and distributions of scores in Figure 4.2.

Table 4.3

Descriptions and Descriptive Statistics for the Scales Based on Items from Student Questionnaire. (Means for Italian, Physical Education Studies and Engineering Added for Comparison).

	N	Min	Max	Mean	Std. D	Cronbach Alpha	Italian	Physical Ed. Studies	Engineering
eAssess	110	1.4	4.0	3.2	0.4	0.85	2.6	2.9	3.2
eAssessP	108	1.2	4.0	3.2	0.4	0.89	*	*	*
Apply	105	1.4	3.0	2.4	0.4	0.34	2.3	2.0	2.2
Attitude	105	1.4	3.0	2.6	0.3	0.45	2.5	2.4	2.6
Confidence	105	1.0	3.0	2.7	0.4	0.78	2.7	2.5	2.7
Skills	105	1.0	4.0	3.3	0.5	0.98	2.9	3.1	3.0
SCUse	105	0.0	334	95.9	62.1	N/A	23.2	18.0	34.2

*Only one form of assessment was used.

For purposes of comparison, means for three other samples of students involved in the larger study of which the current study was a part are presented. These students were assessed by a performance examination only. The reliability of each scale was checked using Cronbach's Alpha coefficient as also reported in Table 4.3. The *eAssess*, *eAssesP*, *Confidence* and *Skills* scales all had acceptable levels of reliability.

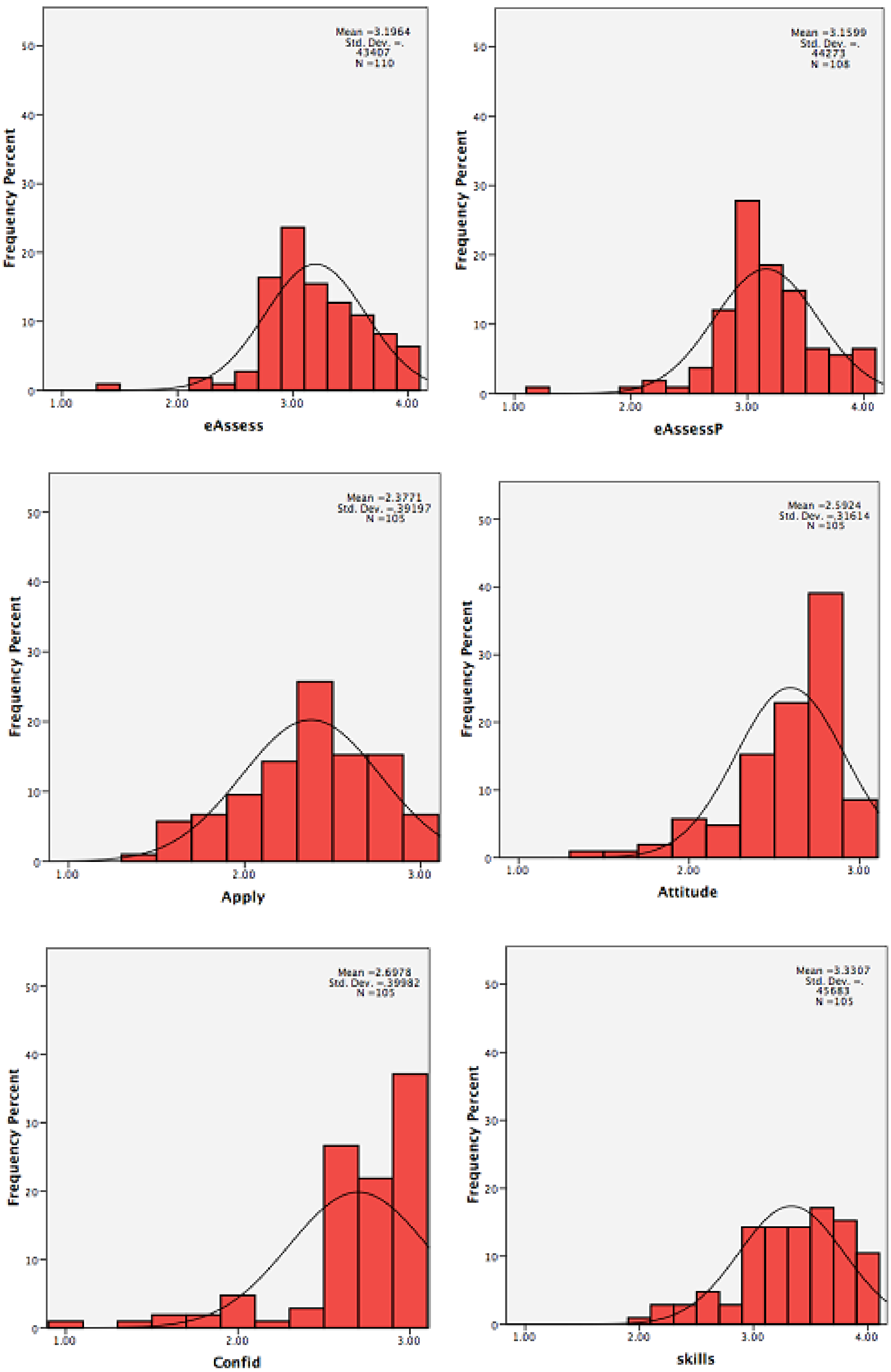


Figure 4.2 Graphs for the distribution of scores for the scales on the student questionnaire.

Though students indicated in item E1(a) that they had little experience in doing examinations on computers (44% indicated no experience), 62% (item E1(b)) felt they would need little or no time to get used to the process. Most of the rest felt they would need some time to do so. For the portfolio (item P1(a)) only 17% indicated that they had no previous experience of completing a digital portfolio and 50% indicated in response to item P1(b) that they would require little or no time to get used to it. Almost all students indicated by response to items E2(a)-(k) that doing the examination on the computer was quick, easy and preferable to the traditional pen and paper examination. Only for development of design ideas, sub item E2(c), did preference fall below 70%, this reflecting the fact that students designed almost exclusively on paper for the examination. Responses to questions P2(a)-(k), which surveyed students' attitude towards completing assessment by portfolio, were overwhelmingly positive with all questions recording more than 80% agreement. Students responded positively to statements describing the authenticity, ease and efficacy of assessment by computer-based examination (*eAssess* mean=3.20) and digital portfolio (*eAssessP* mean=3.16) on a scale of 1 (strongly disagree) to 4 (strongly agree).

In parallel research, students in Italian Studies, Physical Education Studies (PES) and Engineering Studies responded to these items in relation to an examination only, there being no portfolio for these subjects, and the means for these courses have been added for comparison. Students were very positively disposed to completing these types of assessments on computers and felt they would need little time to become accustomed to the process. Compared with the other courses, on average their perceptions were on the more positive side. Overall, the two distributions (*eAssess* and *eAssessP*) were positively skewed with almost none below the midpoint (Figure 4.2).

As might be expected from students of AIT, responses to questions about their disposition towards computers in general, for example whether or not they thought computers were *good for the world*, were positive. This is evident in the mean of 2.6 on the *attitude* scale of 1 to 3. Responses from students in Italian Studies, PES and Engineering Studies were similarly positive ranging from 2.4 to 2.6.Students appeared to be less enthusiastic users of computers for day to day tasks (*Apply* scale mean score 2.4) though this might have been due to the options presented in this question. For example, keeping a list of telephone numbers and addresses (item 10a) would most likely be done using a mobile telephone and not a computer even by those who used computers extensively. Sending a letter to every club member or friend (item 10e) might elicit a negative response for various reasons not least of which are the alternative methods of communication possible.

Most students felt confident around computers, liked using them and felt they were good with them. This is evident in the high mean of 2.7 on a scale of 1 to 3 on the *Confidence* scale (Table 4.3) and in Figure 4.2, where the distribution is highly skewed towards the positive. Student skills were self-assessed by a series of questions as set out in Appendix H. Overall students indicated a high self-assessment of their computer skills; Figure 4.2 shows scale scores highly skewed towards the positive and Table 4.3 a mean of 3.33 on a scale of 1 to 4. Of the types of computer software listed, only in the areas of web authoring and databases did students feel their skills lacked proficiency and their responses, as well as being a measure of ability are to some extent an additional aspect of confidence.

Nearly all the students had home access to the technologies listed in item 5 with mobile ‘phone and MP3 player ownership both more than 90%. Two thirds of the students owned their own laptop computer and 95% had a home broadband Internet connection. At school they estimated that they used computers for an average of 95 minutes per day. These results were consistent across schools irrespective of type or socio-economic background of students.

Interviews with Students

A student forum of four to six students was assembled at the conclusion of each of the practical examinations. Students were prompted to reflect on the portfolio and examination according to seven questions in a semi-structured interview. Supplementary questions were sometimes added to draw out or clarify responses. A summary of results of an analysis of this data across the seven schools is now provided.

Q1 What did you think of the task(s) you were asked to do?

In general across the classes, students were positive about the practical examination as a means of assessment and considered the tasks to be suitable and appropriate though many said the time allocated could have been greater. This is illustrated in the following quotations from the interviews.

“The tasks were well explained and easy to complete” (ZA).

“Pretty easy…I thought the [reflective questions] was a bit confusing sometimes” (WA).

“The last part…about the wording…” (WA).

“Pretty good because you are put under time pressure like a real life situation” (ZA).

“Quite long but quite simple to do” (CA).

“Weren’t difficult but just the time…” (CA).

“They were good tasks but the time was too short” (CA).

“Simple to understand and easy to follow…summed up what we have done this year” (XA).

“More modern than writing things” (RA).

Q2a Were you able to do your best quality of work?

Generally students agreed that they were able to demonstrate their skills and produce good quality work but often qualified this with a comment about time allocation, or for a few, software constraints.

“You can always do better if you’ve got more time” (XA).

“Need more time for planning and producing” (WA).

“I spent too much time on the design.” (WA).

“With the time constraints I don’t think you could.” (ZA).

“With more variety of software we’d have been able to do better stuff” (ZA).

Q2b Did the IT help?

The use of computer technology was cited by all as enabling both in speed, organisation and creativity, as illustrated in the following very typical quotations.

“Yes…definitely easier” (CA).

“It made the design part easier for me as I can’t draw well” (WA).

“Definitely easier to complete [the reflective questions]” (WA).

Q3 How much different was this to how it used to be done?

The contrast to traditional paper based assessment was obvious to all and the preference for this alternative method was unanimous.

“This is a lot better …a lot simpler because you can type with no messy papers if you

make a mistake you can backspace don't have to worry about crossing out" (CA).
"You just go bang I want it there and it's done" (CA).
"This is much better… a lot simpler… practicals let you show what you can do I'm not great at theory" (MA).
"A good change from just writing it up so you had more opportunities to show your skills" (ZA).

Q4 What, if anything, would you like changed in future?

There was a general feeling amongst all groups that more time should be allocated to the practical section of the examination. Some of the other comments are represented in the following quotations.
"Give us example time but let us choose" (CA).
"People know how they work best" (CA).
"Get rid of the audio section even though it didn't work" (CA).
"More time or just make it less to do" (MA).

Q5 Were there any technical problems with doing the activity?

The audio recording was a problem at three of the seven schools and had to be abandoned. Other technical problems mentioned were corrupted files on the USB drives (this occurred on three occasions at MA), student computers at two schools not having software installed (XA two computers and WA about half the computers forcing the student cohort to be split into two halves with half moving to an adjacent computer laboratory). Five computers across all schools stopped responding during the examination and had to be restarted. Two of these were at CA, two at RA and one at XA; the disruption was minimal in all instances. These incidents are reflected in the following students' comments.
"That's my only downside using the computers … if something screws up you don't have anything else" (CA).
"I had to shut down programs three or four times… my page closed before I'd saved" (CA).
"Most of the programs at the start wouldn't even download" (WA).
"Needs to be all set up beforehand to make sure" (WA).

Q6 Were there any other problems with the activity?

Generally comments made here concerned structure of the examination itself, the wording of the reflective questions and criticisms of the school's hardware. A sample of typical quotations is given below:

"Having new computers… it took a while to load applications…" (CA).

"Less reflective questions…" (CA).

"Having faster computers it took a while to load some programs" (CA).

"And some of the questions were really awkwardly worded... like what was your conductivity (sic) and does that mean… it took me five minutes to figure out" (XA).

"The evaluation at the end it just seemed like you asked the same questions four times… I got four words out of a thesaurus copied and pasted those in three or four times" (LA).

Q7 Any other thoughts … or suggestions for developing the use of digital forms of assessment?

Students often reiterated comments from other questions but there were occasional suggestions for improvement, as in the following quotations.

"The microphone at the beginning…maybe if you had USB headsets instead of the traditional red green and yeah USB drivers just sort it out" (LA).

"Having faster computers" (WA).

"The exam was mostly just design with no emphasis on other subjects like databases and spreadsheets...We've done quite a bit of the technical stuff and not much on design" (WA).

Interviews with Teachers

All seven teachers responded fully to 10 questions sent by email concerning their perceptions of the examination and portfolio tasks. Though sentiment was mainly positive, there were some wide-ranging responses and interpretations of the questions. Each question was analysed separately.

Q1. What did you think of the task?

Six of the seven teachers rated the tasks from 'appropriate' to 'excellent' and often qualified this with an explanation of the implementation process. It was here that some light was shed on the reasons for variation or non-compliance with the project guidelines which intended the assessment tasks to form the basis of course assessment.

"The assessment tasks were really good. There was a problem with the compiling of what my students did into the written Design Process document" (CA).
"I needed to follow the framework of the sample exam provided by the Curriculum Council, so the evaluation students submitted to me differed to that submitted to ECU" (LA).
"I thought the tasks were excellent, easy enough for less able students to achieve an adequate result but open ended enough for the good students to excel" (WA).
"I felt that the task was excellent-There was a lot of scope in terms of design" (ZA).

Q2. What did you think of the structure of the activity?

Reactions to the structure of the tasks was mixed and again accompanied by some explanation of what was actually done in contrast to what had been intended. There was some criticism of the time allocation for both the portfolio and the examination. The rigidity of the timing of the examination was perceived negatively by three of the seven teachers. Three teachers were positive about the structure and the timing, as shown in the following quotations.

> The timing fitted in exceptionally well with the Course Outline and students were able to follow the instructions. The only problem I noted after the conclusion of the task was that students had been asked to submit the two artefacts and write ups and several students did not comply with this request (LA).
>
> The structure of the activities was good. The timing and structure was fine with the portfolio and instructions were quite clear. There were some timing issues with the exam. Students seemed rushed. The response section of the practical exam, students wrote a response which was marked and then recorded the written response which was marked separately. Not sure if that was the intention but the marking ended up being a lot for that small part of the work (MA).
>
> It seemed inappropriate to suddenly interrupt them and tell them they now had to do something else, like plan or record a response. It seemed unfair not to let them go back and add to work done is part A after the time for that portion of the exam had expired. Design is an iterative process, and the ability to review and improve your work is a work habit to encourage (WA).

Q3. What were the students' reactions to the activity?

The teachers' perceptions of the reactions of the students were expressed in positive terms by all teachers with the exception of WA. Some added their own interpretation of their students' wishes and desires.

"They were happy enough with the practical but I got a feeling they weren't all that keen about a practical exam" (CA).

"Positive, although they felt under pressure, and we had real problems with the school network playing havoc with their files" (ZA).

"The students enjoyed the portfolio task. Students liked the idea of the practical exam (MA).

They were unhappy with the marks distribution. They said the exam was easy, but most performed surprisingly poorly" (WA).

"The exam was OK, although creating a brochure was not what we really wanted to do" (ZA).

Q4. What do you think of its potential?

In response, the majority of teachers were positive about the idea but almost all had reservations about the structure and implementation. Two mentioned concerns regarding hardware. Three teachers indicated that a portfolio *and* an examination should form the basis of the assessment. There were suggestions that an externally assessed portfolio could replace the moderation of students' coursework. One teacher (WA) highlighted the fact that students could clearly see what others were doing and this invalidated the examination as a test of an individual's ability. The range of responses is represented in the following quotations: "After marking the TEE (Tertiary Entrance Examination) AIT exam, a practical exam for the production component couldn't come quickly enough. If the external marking of a portfolio does away with the moderation process, I'm all for it" (CA).

> The potential of a practical exam is good however it would need to be a little more flexible such as a website/pages, logo, brochure etc. as the range of practical assessments in the course varies a lot. The portfolio I can see will be very time consuming for students to create however it would be a better way of assessing than just a paper exam (MA).

"There is great potential for AIT and Engineering Studies. As we proceed further down this path, the process will improve. My students are telling me this should have happened earlier (RA)". "There is way too much potential for seeing what other people are doing and get ideas from others. Any student with the intention of cheating could have done so too easily (WA)".

Q5. What did you think of the quality of work produced by your students for this task?

The quality issue drew a variety of interpretations and responses with teachers evaluating student work in broad and relative terms, for example making comparisons to previous years. Two teachers (MA & WA) described their students work as poor or basic and put this down to inexperience or lack of choice of product in the examination. "I felt I got better results from the students this year as opposed to last year because of the wording of the tasks supplied (CA). "The quality of some of the portfolios were good and in some cases better than expected. The quality of the exam I thought was quite poor given that it was practical and not paper-based (MA)." "Students produced basic work in exam. It would have been better to have a choice of ICT products (ZA)."

Q6. Were you surprised by the performance/attitude of any students?

There were mixed interpretations and responses to this question and nothing exceptional, either positive or negative, was reported. In general the performance of students on the examination was perceived as underachieving (two comments) while performance on the portfolio was neutral or better than expected. These responses are illustrated in the following quotations: "I was disappointed with the students' exam (MA)." "I was pleased with the quality of most students portfolio work" (CA).

Q7. What was the general feedback from students?

Responses were again mixed with three of the seven teachers claiming students were happy with the tasks and positive about more of the same. There were some negative comments again about timing; the lack of time available to manage the portfolio and complete the course, and the wording of the reflective questions. The typical mix of responses is represented by the following quotations: "They were happy enough with the tasks. They dragged their feet with the process document. I'm not sure about the exam" (CA). "Happy with the exams and would like to see more of the same" (RA). "They thought the practical exam was harder to complete than they had thought" (XA). "Unhappy that they did not have enough time-preference for a choice of products-computers too slow-desire for a more simplified set of task instructions" (ZA).

Q8. Were there any technical problems with implementing the activity?

Hardware failure and underperformance predominated in the responses to this question. Though the audio recording failed partially at three of the seven schools only two teachers mentioned it here. Other technical problems cited were corruptly imaged USB flash memory (MA & WA teachers) and general complaints about speed of computers and school network issues (RA & XA).Two teachers reported no technical problems. “Many of the USB files were corrupt. One student had to try 3 thumb drives before he got all the files onto his computer” (WA). “Just school computers and downtime which affected us severely” (ZA). “Very little technical problems” (MA).

Q9. Were there any other problems with implementing the activity?

Four teachers reported no other issues with implementation. Of the other three, two mentioned time pressure and one the perceived repetition of reflective questions in the examination.

Q10. Any other thoughts or suggestions for developing the use of digital forms of assessment?

Teachers responded fully to this question with a wide variety of suggestions for improvement. Three teachers were dissatisfied with the example marking keys provided and two said they had written and used their own. This might explain the absence of consistency between external assessors and teachers which is evidenced later in this chapter. There were some sound suggestions concerning integration of activities across the entire syllabus, for example presenting some of the theoretical aspects as websites. There was also the suggestion of including software as well as sample files on the USB flash memory supplied to the students. Some of these suggestions are illustrated in the following quotations:

> Create a bootable USB memory stick and have all the required applications and programs for the exam so that students do not have to rely on a local network. In this way there is more control over the whole environment (RA).

“A better fit with the entire syllabus as there was too much work to cover the syllabus in its entirety and complete the portfolio and other tasks as well” (MA). “A better, more comprehensive marking scheme” (MA). “The marking key needs a serious revamp” (ZA).

Teachers' comments were candid and unreserved, spanning the full spectrum from glowing praise to scathing criticism. Though unrepresentative, in that they voluntarily signed up to take part in the study, the comments overall were mainly positive about the implementation of both forms of assessment. In spite of their enthusiasm, the comments of teachers and the fact that many pursued their own assessment strategy suggests that they would require a little more convincing before adopting any new assessment methods.

Marking of Student Work

Each student's work was marked by external assessors and his or her teacher. The former marked all student work (portfolio and examination) analytically and the production component of the examination for 60 selected students using a *comparative-pairs* method. Students' work was uploaded to an online repository of files stored on a server and arranged in folders assigned to each student. A web-enabled, password protected, database management system using the *Filemaker Pro* software allowed the content of each student folder to be displayed within the online marking tool.

Analytical Marking Tool

The analytical marking tool was developed for the main project using *Filemaker Pro* software with embedded rubrics displaying the marking criteria for the portfolio and examination tasks as set out in Appendixes D and E. The tool displayed the students' work on-screen and allowed the judgements of the assessors to be digitally recorded. The development environment allowed the tool to be deployed using a web browser over the Internet and to be password protected. After authentication, the assessor was able to see a list of all students by student ID within each class. For each student there were five buttons which opened the five marking screens, one per component, each with a specific rubric and links to student work as shown in Figure 4.3. The marking tool design was based on a 20 inch screen allowing the assessment criteria to be displayed on the left side and the student work sample on the right for each marking window. The tool incorporated a marking key (rubric) based on the assessment criteria developed for the task. Marks were recorded by selecting the appropriate radio button as shown in Figure 4.4. Hyperlinked buttons also allowed movement between marking windows. Space was allocated for recording of notes and comments as required. The tool was designed to handle all clerical functions such as totalling of marks and saving comments. Analytical marking of students' work was completed entirely using this online marking tool. Both assessors were experienced computing teachers.

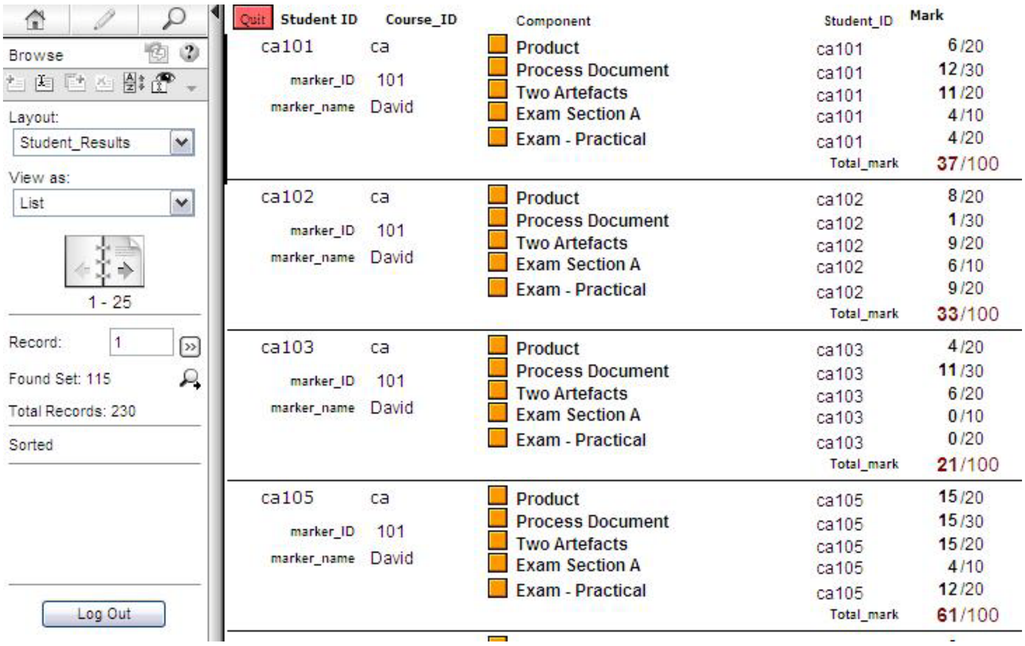


Figure 4.3 The marking tool showing the five assessment components

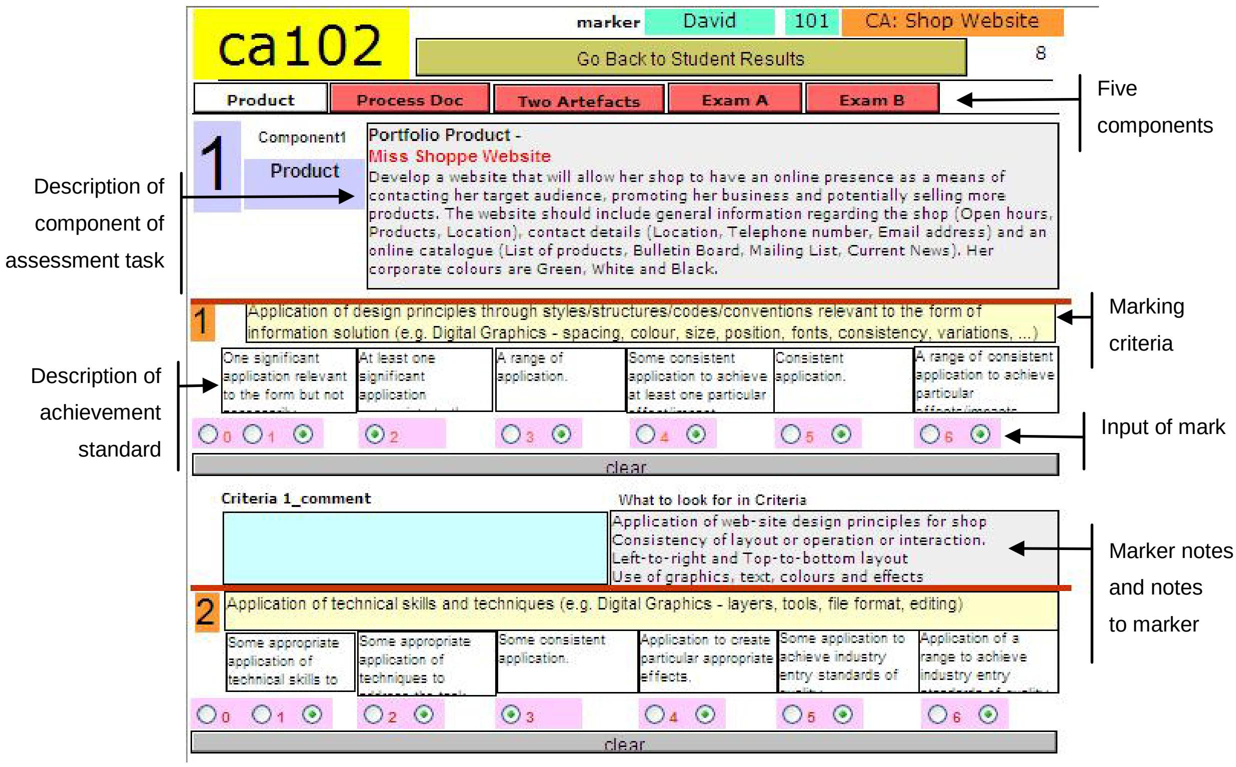


Figure 4.4 The marking tool (left side of screen) showing the marking rubric for component one for CA102

Comparative-pairs Marking Tool

The comparative-pairs marking tool used the same on-line database management system as the analytical marking. The tool was developed for the main project using *FileMaker Pro* and deployed on the Internet with minor modifications. A web enabled database was designed to display samples of two students' work (A and B in Figure 4.5) from the production examination side-by-side, with interactive controls for recording of the marker's choices located between them.

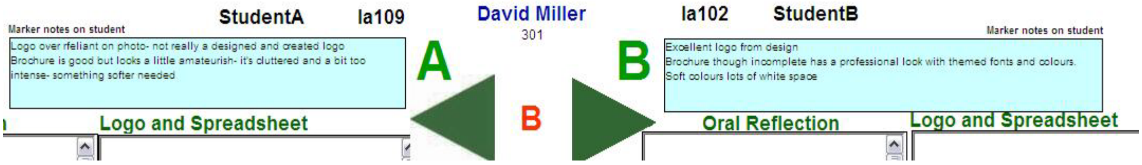


Figure 4.5 Comparative Pairs Marking showing marker notes and holistic selection of student B (la102)

Figure 4.5 shows the marker notes for two students La109 and La102. The assessor was required to make four choices by clicking on large green arrows pointing toward the student they wished to select; three based on specific criteria (Design Process, Technical Proficiency and Design Principles) and one overall or Holistic judgment. Figure 4.6 (alongside) shows the selection of *student B* over *student A* according to all four criteria. A short description of what to consider for each of the choices was provided. A text field was available for each student so that assessors could record their comments on the students' work and have these re-appear each time that students' work was involved in further comparative assessment. When completed, assessors clicked on a button to bring up the next pair.

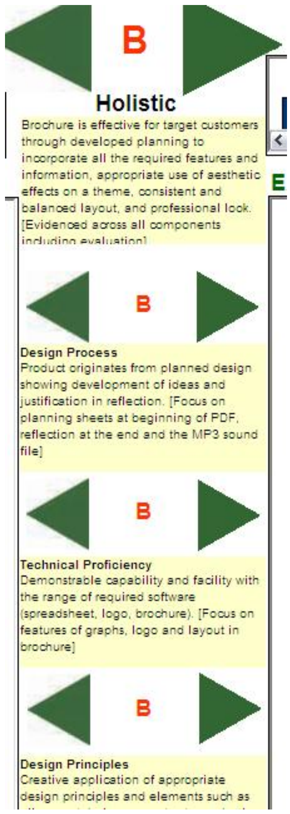


Figure 4.6 Comparative pairs marking: selection based on three criteria.

Interviews with assessors

Assessors reported that analytic assessment using online marking tools had several advantages over paper based systems. Firstly, marking could be done from any location either inside or outside Australia and at any time of day. Totalling of scores was automated and this improved reliability and speed. The main drawbacks reported were viewing of certain types of files which could not be displayed in a web browser. These were often large and had to be downloaded prior to opening. This was often time consuming particularly from outside Australia in countries where download speeds were low. Failure to display a file successfully compromised marking accuracy as such files were treated in the same way as those which were either missing or corrupted and scored as zero.

The comparative-pairs marking, involving the work of 60 selected students on the practical component of the examination had similar advantages to the analytical marking. Speed of marking, though slow initially, increased as familiarity with the system and marking criteria increased. Apart from the extended download times, no difficulties were reported using the system for either type of marking.

Marking by Teachers

Teachers were requested to assess their students' work on the portfolio and examination using their own marking criteria and to forward their results together with an overall student semester mark. A sample rubric was provided based on the analytical marking tool. In the event, only two complete sets of marks were received (MA and ZA) with four others giving marks for one or more components. Various reasons were given for this with the fundamental misunderstanding of how the study was intended to work being widely prevalent. It came to light, in discussions with teachers during school visits or later during teacher interviews, that two of the seven schools (LA, ZA) were running the portfolio and examination as additional tasks and not as integral parts of the assessment scheme for the semester. For example, from the teacher interview (LA), "They had no problems accepting the situation of doing an exam that would not be credited as part of their AIT Unit 2B…"

Results of Marking

Students' work from the portfolio and the examination (components one to five) was assessed analytically by two external assessors using the online analytical marking tool. The examination component five of a sample of 60 students was assessed again by five assessors using a comparative pairs' methodology. Further, it was intended that all student work be marked independently by the class teacher using their own marking system. These results, together with an overall semester mark were requested for analysis. This section contains detailed analysis of:

- the marks awarded by the external assessors for the five components of the study;
- a comparison and analysis of the marks awarded by teachers for the examination and overall semester mark with the marks awarded by the assessors;
- a comparative analysis of the portfolio and examination using assessors and teachers' marks;
- an analysis of the comparative-pairs marking;
- a comparison of analytical and comparative-pairs marking; and
- analysis of the portfolio and examination as assessment instruments.

Marking by Assessors: Analytic

Two external assessors, both very experienced senior computing teachers, worked entirely independently and solely using the online analytical marking tool. All components of the portfolio and the examination were marked with scores totalled and no allowance being made for missing, unavailable or corrupted work. For each school, the total scores of marking were compiled into a table showing the scores for each individual student using the headings shown in Figure 4.7.

	Assessors marking (Total)			Assessors (Average)		Teacher	Rank	
St ID	Ass1 (%)	Ass2 (%)	Ave (%)	Pfolio (70)	Exam (30)	Sem (%)	Ass Ave	Tch Sem
ca101	37	30	33.5	23.0	10.5	70	63.0	3.0

Figure 4.7 Structure of the table of results for each class.

The rank for the average of the external assessors marks (*Ass Ave*) was for all 115 students whereas the teacher's semester mark (*Tch Sem*) rank was just for the class. The rank was the position out of the whole sample of 115 students spread over seven schools. For example, an average assessor rank of 86.5 meant that this students total ranked 86th equal out of 115. The teacher's semester rank was the position of the mark awarded by the teacher compared to the marks awarded by *all* other teachers in the sample.

Table 4.4 shows descriptive statistics for the scores from analytical marking for each school. The scores were derived by averaging the marks awarded by the two assessors. Assessor mean scores for component 1, the portfolio product and component 5 the examination practical section are shown separately.

Table 4.4

Descriptive Statistics of Average Scores Awarded by Two External Assessors for all Students.

		All Tasks			Portfolio Components 1-3			Exam Components 4-5			Component 5 (2h practical)			Component 1 (Digital product)		
Class	N	Range	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range	Mean	SD
CA	20	14-67	32.2	13.7	0-53	20.6	12.3	0-18	11.6	4.3	0-12	7.8	3.4	0-15	7.3	4.6
LA	10	18-56	40.6	11.2	12-40	27.0	8.8	0-23	13.6	6.5	0-16	8.8	4.6	3-14	8.6	3.2
MA	12	25-69	47.4	13.6	13-47	32.3	10.5	10-23	15.1	4.0	6-15	11.3	2.4	5-17	10.4	3.6
RA	14	13-63	31.2	14.8	0-43	17.4	12.4	9-21	13.8	3.6	6-15	9.9	2.5	0-14	7.4	4.7
WA	14	8-17	12.6	2.9	NA	NA	NA	6-17	12.6	2.9	3-13	9.0	2.5	NA	NA	NA
XA	29	17-66	39.8	13.5	0-47	23.4	11.7	8-24	16.4	3.7	5-16	11.5	2.7	0-17	6.6	6.5
ZA	16	26-75	53.6	15.1	14-53	35.8	11.5	9-25	17.8	4.4	7-17	11.9	2.4	0-17	11.1	4.5
All	115	8-75	36.9	17.2	0-53	22.3	14.6	0-25	14.6	4.6	0-17	10.1	3.2	0-17	7.2	5.6

The mean of the scores awarded to each school for all tasks allowed a comparison between schools to be made. The mean for WA is an outlier resulting from the non-submission of the portfolios. The two schools with the highest mean scores were MA (Independent school) and ZA (Catholic School); the two with the lowest mean scores (ignoring WA) were the state schools CA and RA. Average performance on the examination followed a similar pattern, as measured by mean of scores awarded, but was much closer with ZA and XA (both Catholic) followed by MA (Independent). The mean scores concealed the fact that all schools with the exception of LA (ignoring WA) had some high performing student as indicated by the upper limit of the range. The spread of scores for all schools again with the exception of LA (ignoring WA) were similar with standard deviations between 13.6 and 15.1 marks as shown in Table 4.4.

The correlation between total scores awarded by each of the two assessors was analysed for each class with the resulting coefficients shown in Table 4.5. The larger samples (20 or more students) appear to exhibit stronger correlations between assessors. RA was an exception, where the sample size was small and the average mark ranked sixth out of seven suggesting a poor average quality of work.

Table 4.5

Correlation Coefficient Between Total Scores (Portfolio Components 1-3 and Examination Components 4-5) of Two External Assessors by School

Case	Number of Students	Correlation coefficient between assessors
CA	20	0.88**
LA	10	0.55
MA	12	0.75**
RA	14	0.90**
WA	14	0.66*
XA	29	0.91**
ZA	16	0.78**

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

For class RA 28 out of 140 assessed components (2 assessors x 5 components x 14 students) were missing or scored zero. As a result it is not surprising that the correlation coefficient was high. Apart from the small sample size, weak correlations between assessors for LA may be due to difficulties one assessor had in accessing component three of the portfolio, the two additional digital artefacts. In five of the ten instances, assessor 1 scored a zero whereas assessor 2 recorded a mark. This discrepancy is most likely due to the fact that assessor 1 was unable to locate or open the work samples and hence treated them as corrupted or missing files.

The results of marking were analysed in terms of the components of the assessment task. Some of the descriptive statistics resulting from this are shown in Table 4.6. Calculation of the inter-assessor correlation coefficient, a measure of the reliability of the scores awarded, revealed a high and significant correlation between the total marks awarded by the two assessors ($r=0.89$, $p<0.01$). Though there were minor discrepancies in the partial and total scores awarded by each assessor, the correlation coefficients between assessors for scores awarded for each of the five components were also moderate to high, and significant. This indicates good internal consistency across all components pointing to the precision of the analytical marking tool and to its consistent interpretation by the assessors. The marks awarded for the portfolio are more strongly correlated than those awarded for the examination. However, the suggested greater strength of the correlations for the portfolio components may be due to the fact that many portfolios were incomplete and attracted component marks of zero from both assessors (giving rise to unitary correlation) whereas the examination was completed by all and some marker variation might be expected. The effect of missing components of the portfolio was investigated.

Table 4.6

Descriptive Statistics from Analytical Assessment of 115 Students by Component and by Assessor.

N=115		Assessor	Range	Mean	SD	Correlation Between assessors
Portfolio	Component 1 Digital product	Ass 1	0-18	5.7	6.8	0.84**
		Ass 2	0-19	5.9	7.6	
	Component 2 Design Document	Ass 1	0-28	7.8	8.7	0.90**
		Ass 2	0-29	8.6	9.4	
	Component 3 Two further artefacts	Ass 1	0-18	5.5	6.6	0.82**
		Ass 2	0-19	5.0	5.8	
Examination	Component 4 One hour theory	Ass 1	0-8	2.1	4.1	0.60**
		Ass 2	0-10	2.2	4.8	
	Component 5 Two hour practical	Ass 1	0-20	3.6	10.9	0.62**
		Ass 2	0-18	3.3	9.6	
	Total	Ass 1	7-83	17.0	36.6	0.89**
		Ass 2	9-80	18.2	37.1	

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

Table 4.7 shows the percentage of zero marks awarded for each component by the two assessors. Almost 30% of portfolio components (204 out of 690) were assessed as zero or were missing, whereas less than 3% of examination components (17 out of 690) were assessed as zero or were missing.

Table 4.7

Percentage of Missing or Zero Marks Awarded for Components of Portfolio and Examination by Two External Assessors

	Portfolio			Examination	
	component 1	component 2	component 3	component 4	component 5
Assessor 1	30%	30%	29%	9%	1%
Assessor 2	29%	31%	30%	3%	3%

The correlation coefficients between assessors were recalculated with the scores of all students awarded zero by both assessors removed. These are presented in Table 4.8. The correlation coefficients are weaker but still significant ($p<0.01$). The correlation coefficients, though still moderate and significant for all components, are markedly weaker for the portfolio but less so for the examination results. It should be noted that in three instances for component one and five instances for component three a score was awarded by one assessor and a zero was awarded by the other. This is explained by the failure to correctly download and open the work for marking, these components often being large files. The relatively high correlation for component 2 (the design process document) may be explained by the more structured nature of this component affording less interpretation in marking.

Table 4.8

Correlation Between Two External Assessors on Total Marks Awarded for Components of Portfolio and Examination With Any Pairs of Marks of Zero Awarded Removed

	Component	Number	Pearson Correlation (paired zeros *removed*)	Pearson Correlation (paired zeros *included*)
Portfolio	component 1	82	0.59**	0.84**
	component 2	82	0.78**	0.90**
	component 3	83	0.52**	0.82**
Examination	component 4	112	0.54**	0.60**
	component 5	113	0.55**	0.62**

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

Marking by Teachers: Analytic

Teachers were requested to mark the examination separately (both theory and practical sections) and award a semester mark for the student which included the portfolio. The aim of the study was to have teachers mark the *same* content as the external assessors using their *own* assessment methods and to forward this together with the overall semester mark. However, there was widespread misunderstanding of this intention and the content actually marked by the teachers and the proportion included in the final semester mark, varied according to school and was not always obvious to the researcher. Of the 115 students, only 58 final semester and 26 examination marks were received from teachers. Further, student portfolios were submitted in various stages of completeness. Table 4.9 shows a summary of the descriptive statistics for the scores received. When comparing scores awarded by teachers with those awarded by external assessors the following should be noted:

i) No separate teacher scores for the portfolio were received
ii) Two out of seven teachers supplied marks for the examination but in one case without detail of whether or not the whole examination (including the reflective questions) was assessed. Only in one case was a full breakdown of marks supplied.
iii) Three out of seven schools supplied a semester mark for each student; however the makeup of this mark was unknown to the researcher.

Table 4.9

Descriptive Statistics of Results Supplied by Teachers from Marking Examination and Semester Mark.

	N	Minimum	Maximum	Mean	Std. Deviation
Teacher examination %	26	28.1	91.8	57.9	17.1
Teacher semester %	58	9.0	84.0	56.5	18.3

With due regard to these variations, comparison between assessor and teacher marks is presented in Table 4.10 (scores) and Table 4.11 (ranks). There were moderate and significant correlations between Assessor Average and Teacher Semester mark ($r=0.62$, $p<0.01$) despite the fact that the content assessed may have been somewhat different.

Table 4.10

Correlation Between Marks Supplied by Teachers for Examination and Semester and Marks Awarded by Two Assessors

	Assessor 1 Total	Assessor 2 Total	Assessor Average	Teacher Examination %	Teacher Semester %
Assessor 1 Total	1.00	0.89**	0.97**	0.25	0.60**
Assessor 2 Total		1.00	0.98**	0.36	0.58**
Assessor Average			1.00	0.32	0.62**
Teacher Examination %				1.00	0.93**
Teacher Semester %					1.00

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

This suggests that either the quality of students' work was recognised by both external assessors and teachers alike or that what was assessed constituted a similar measure of student capability. For example schools ZA and LA did not include the portfolio or examination mark as part of their semester mark. These activities were done as supplementary tasks. In both cases, preparation for external examinations was cited as the reason for this, with students working towards an external, written paper.

Table 4.11

Correlation Between Ranks of Marks Supplied by Teachers for Examination and Semester and Rank Of Marks Awarded by the Two Assessors Using Analytical Marking

Rank of	Assessor 1 Total	Assessor 2 Total	Assessor Average	Teacher Examination %	Teacher Semester %
Assessor 1 Total	1.00	0.91**	0.97**	0.58**	0.24
Assessor 2 Total		1.00	0.98**	0.57**	0.35
Assessor Average			1.00	0.61**	0.31
Teacher Examination %				1.00	0.94**
Teacher Semester %					1.00

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

Where available, teachers' examination and semester marks show very strong and significant correlations ($r=0.94$, $p<0.01$) indicating that students who did well in the examination also did well over the whole semester. Without knowing exactly how the semester mark was derived and what weighting the examination contributed to it, it is not possible to say for certain why this was the case. It is possible that the examination was measuring the same aspects of student performance as the semester coursework and that the teacher's in depth knowledge of the student's ability had a bearing on the assessment. Obviously the latter was not a factor for the external assessors and it should be noted that correlation between external assessors' total scores (for the portfolio and the examination) are very weakly correlated with teachers' semester scores.

Comparative-Pairs Marking and Analysis

Comparative-pairs marking involved assessment of the production examination for a reduced sample of 60 students. These students were chosen because their practical work samples were equivalent in the degree of completeness and had no missing sections. In particular they all had an audio response file for the examination. Five assessors each completed the pre-determined set of comparisons between students using a digital marking tool. All five were computing teacher experts, two having being involved in the analytical marking. One holistic and three specific assessment criteria were developed for the comparative-pairs marking from the criteria previously developed for the task. These criteria were:

Holistic Criterion: Brochure is effective for target customers through developed planning to incorporate all the required features and information, appropriate use of aesthetic effects on a theme, consistent and balanced layout, and professional look. [Evidenced across all components including evaluation]

Specific Criterion 1: Design Process. Product originates from planned design showing development of ideas and justification in reflection. [Focus on planning sheets at beginning of PDF, reflection at the end and the MP3 sound file]

Specific Criterion 2: Technical Proficiency. Demonstrable capability and facility with the range of required software (spreadsheet, logo, brochure). [Focus on features of graphs, logo and layout in brochure]

Specific Criterion 3: Design Principles. Creative application of appropriate design principles and elements such as alignment, balance, contrast, emphasis, harmony, proportion, proximity, repetition, unity, and white space. [Focus on brochure and logo].

Analysis of Results from the Comparative-Pairs Marking

Rasch Uni-dimensional Measurement Model (RUMM) software (RUMM Laboratory, 2011) was employed to analyse the results of the multiple comparisons of the 60 selected production examination exemplars. A sample of the resulting output of this analysis is shown in Table 4.12. Each of the 60 exemplars was allocated an identifying 'Code' (column 1). *Preferred* (column 2) is the actual number of times the exemplar was preferred in all the *Involved* (column 3) comparisons. The exemplars were ranked from best (most number of times preferred) to worst (least number of times preferred). *Estimate* (column 4) is the exemplar location in logits (logarithmic units of measurement) and is the determinant of the rank order of the exemplars. *Std Err* (column 5) is the standard error of measurement. *Outfit* (column 6) is an index of whether the pattern was more or less Guttmann like, that is whether the exemplar was consistently rated. It is expected to have a value of about 1.00.

Table 4.12

Sample of Output from RUMM Analysis Showing Headings and the First Three Records

Code	Preferred	Involved	Estimate	Std Err	Outfit	Chi Sqr	Degrees Freedom
40	56	59	3.66	0.63	0.82	46.56	57.03
35	55	59	3.34	0.56	1.18	67.26	57.03
43	53	59	2.83	0.48	0.69	39.24	57.03
etc.							

A Separation Index (SI) was calculated as an indicator as to whether or not the scores for the exemplars were sufficiently diverse in quality to assure a broad enough range for the purposes of comparison. The SI is calculated as a number from 0 to 1, with values closer to 1.00 being more desirable. If the value is close to 0.00 (up to about 0.3 or 0.4) the range is too narrow. If it is above about 0.7 the separation is reasonable and if it is above 0.8, the separation is good. Inter-rater reliability analysis was undertaken in order to assess an individual judge's consistency with the judgments of the other judges in the group. This *Outfit Statistic* should in this instance be between 0.5 and 1.5.

The group reliability is defined as the average of the individual rater reliability indices. Table 4.13 below summarises the results of the analysis. Detailed results appear in the Appendix M. The results show that the scores for the exemplars were sufficiently disparate to be reliably compared according to the criteria. Further, the inter-rater reliability, the extent to which assessors obtain the same result when using the assessment criteria, was also good.

Table 4.13

Separation Indices and Inter-Rater Reliability Coefficients for AIT

Type of Judgement	Separation Index	Intra-rater Reliability Coefficient					
		Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Overall
Criterion 1	0.94	1.06	0.80	1.26	1.15	1.01	1.05
Criterion 2	0.95	0.88	0.69	1.02	1.43	1.09	1.02
Criterion 3	0.95	1.34	0.83	0.90	2.52	0.97	1.31
Holistic	0.96	0.91	0.60	1.10	1.02	1.37	1.01

Comparison of Methods of Marking: Analytical v Pairs

Only 60 exemplars of component 5, the practical section of the examination, were marked with *both* analytic and comparative-pairs marking. Correlations between the two methods of assessment for this component for these students are shown in Table 4.14. A strong and significant correlation ($r=0.73$, $p<0.01$) was found between the scores generated by the two methods of marking. Correlations between teacher examination scores and comparative-pairs scores are low but it should be noted that teacher examination mark data was received for only 27 of these 60 students. Correlations between scores awarded by each assessor, though significant, were much weaker than for the whole sample (r=0.43, p<0.01). Possible reasons for this are difficulty or error in applying the marking rubric and the subjective nature of many of the judgements required.

Table 4.14

Correlations Between Marking Methods for Practical Component of Examination (Component 5) Only.

	Ass Average	Assessor 1	Assessor 2	Pairs Holistic	Tch_Exam	Tch_Sem%
Assessor Av.	1.00 (60)	0.86** (60)	0.83** (60)	0.73** (60)	0.17 (29)	0.39* (60)
Assessor 1		1.00 (60)	0.43** (60)	0.69** (60)	0.11 (29)	0.42* (27)
Assessor 2			1.00 (60)	0.55** (60)	0.18 (29)	0.20 (27)
Pairs Holistic				1.00 (60)	0.33 (29)	0.47* (27)
Tch_Exam					1.00 (29)	0.82* (08)
Tch_Sem%						1.00 (27)

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

Table 4.15 shows the correlations between the scores for each criterion of the pairs-comparison marking of the production component of the examination and the average of the analytical marking of this component by the two assessors using analytical marking. *Pairs Holistic* is the derived overall score awarded to the student's examination work based on the comparisons it underwent by the five judges in the pairs marking using the holistic criteria.

Table 4.15

Correlation Between the Pairs-Comparison Marking of Component 5, the Practical Component of the Examination and the Analytical Marking of the Examination

	Pairs Holistic	Pairs Criterion1	Pairs Criterion2	Pairs Criterion3	Teacher Exam %	Teacher Sem %	Assessor Average
Pairs Holistic	1.00	0.84**	0.92**	0.97**	0.33	0.47*	0.73**
Pairs Criterion1		1.00	0.74**	0.85**	0.18	0.42*	0.62**
Pairs Criterion2			1.00	0.90**	0.46*	0.43*	0.73**
Pairs Criterion3				1.00	0.33	0.46*	0.70**
Teacher Exam %					1.00	0.10	0.17
Teacher Sem %						1.00	0.36**
Assessor Average							1.00

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

Pairs Criterion1, *Pairs Criterion2* and *Pairs Criterion3* are the scores awarded for criterion 1, criterion 2 and criterion 3 of the student's work derived from the pairs comparisons. *Teacher Exam* is the mark awarded by the teacher for the examination (theory and practical) where received. *Teacher Sem* is the teacher's mark for the student for the whole semester's work, where received.

As might be expected, the three criteria (pairs marking) and holistic criterion were highly correlated with the assessors' average score from the analytical marking. Correlations between the teachers' examination scores and the scores for the pairs marking criteria were low, with the exception of criterion 2 ($r=0.46$, $p<0.05$), and the absence of a relationship once again highlights the fact that the exact criteria and method of marking by teachers was not known to the researcher. However, the teachers' semester mark appears to be moderately but significantly correlated with all three criteria in the pairs marking. ($r=0.42$, $p<0.05$; $r=0.43$, $p<0.05$ and $r=0.46$, $p<0.05$ respectively). Although the correlation between analytical marking score and pairs holistic marking score was strong and significant ($r=0.73$, $p<0.01$) it might be expected to show even less variance if the methods are indeed equivalent in their accuracy. Table 4.16 shows the *rankings* of the 60 exemplars marked by both methods.

Table 4.16

Ranking of the Practical Component of the Examination Marked by Analytical and Pairs Methods (N=60)

Stud ID	Rank analytic	Rank pairs	Stud ID	Rank analytic	Rank pairs	Stud ID	Rank analytic	Rank pairs
za101	1	5	xa103	19	33	za103	38	23
la102	2	9	ma109	19	31	wa115	38	20
xa123	2	8	ma101	19	29	za114	38	17
za110	2	3	ma111	19	23	ra110	44	38
ma108	5	43	xa114	19	20	za108	44	17
xa118	5	1	la109	19	16	wa106	46	57
xa116	7	11	ma104	27	51	wa112	46	51
za109	7	9	wa114	27	26	wa109	48	47
za117	9	33	xa107	27	23	wa102	49	56
xa106	9	13	za115	27	19	xa122	49	38
xa108	9	12	za112	27	15	za107	49	27
ma107	12	43	ra104	32	43	ra103	52	47
xa111	12	6	xa110	32	36	la101	52	28
xa126	12	3	xa104	32	31	za104	54	38
xa112	12	2	wa101	35	38	wa113	55	57
za106	16	22	za105	35	37	wa104	56	60
wa107	16	13	za102	35	29	ra101	56	54
za116	16	6	xa121	38	50	ma105	56	53
xa113	19	55	ma103	38	49	wa111	56	46
ra105	19	33	za113	38	42	wa108	60	59

Even though correlation between ranking for each method remains strong and significant ($r=0.72$, $p<0.01$) there are some major discrepancies.
Whilst a difference in rank of a few places might be explicable in terms of random effects and marker error, the greatest difference was 38 places; work sample MA108 was ranked 5th out of 60 by analytical marking and 43rd out of 60 by pairs marking. Without allocating grade boundaries it is not possible to say how many grades separate these positions. What is possible is some further analysis of the work sample in order to attempt to explain the discrepancies. The ten results with the largest difference in ranking are shown in Table 4.17

Table 4.17

Ten Results Showing the Greatest Difference in Ranking of Component 5, the Practical Component of the Examination Marked by Analytical and Pairs Methods (N=10)

Student ID	Rank analytic/60	Rank Pairs/60	Comment
ma108	5	43	Incomplete brochure but good design, original logo and good climatic graphs. Marker error (analytical) on brochure.
za117	9	33	Brochure not complete and would not work as tri-fold. Incorporates original logo. Comprehensive design brief
ma107	12	43	Extensive design brief logo good graphs complete but brochure looks amateurish- poor impression
xa113	19	55	Component parts complete- good logo and graphs but brochure not fit for purpose- incomplete
ma104	27	51	Not a brochure but a series of pages with information added-component parts well completed designs good
za114	38	17	Brochure quite impressive and complete. Nice effects. No logo and limited design work
wa115	38	20	Completed brochure with all parts included well laid out with neat logo-looks a bit amateurish however
za108	44	17	Very limited designs logo just a photo plus text, however completed product shows some flair
za107	49	27	Limited design, reflection. Product though amateurish is complete
la101	52	28	Few design notes no logo (just text) Brochure looks professional nice graphs, complete and good colours

In five of these the analytic marking produced the higher ranking and in the other five the pairs marking gave rise to the higher ranking. The comment field derives from a re-evaluation by the researcher of the examination product for each student. This was achieved by referring back to the on-line marking tool and noting any assessor comments and checking that the marks awarded according to the analytical marking rubric were free from obvious errors and omissions. Each work sample was also viewed from the perspective of the holistic marking criterion to see where the sample met or failed to meet aspects of the criterion statement as set out in Appendix L.

A possible reason for the disparity is the difference in weighting and marking criteria between the two methods. The analytical marking awarded 10 marks out of 20 for the practical components of the brochure itself (creating a logo-4, drawing graphs-2 and creating a brochure-4) the remainder being awarded for design, selection of software, technical proficiency, file formats, reflection, design standards and conventions. Even though the pairs marking criteria allude to these, it is the visual impact of the brochure, whether or not it looks fit for purpose, which informs the holistic judgment and on this turns the perceived superiority of one product over another. In other words the marking criteria though similar, are different in emphasis and weighting. It is possible to produce an excellent brochure without adequate design and analysis and in such cases a disparity between the scores from the two methods is evident. Equally, it is possible to produce excellent design and analysis whilst failing to produce a brochure which has the necessary visual impact. This could be due to poor technical proficiency or insufficient working time. Either of these mismatches appears to lead to a disparity in ranking and further analysis of the components of each assessment method is shown in Table 4.18.

Table 4.18

Ranking by Criterion and Analytical Marking Ranking for the Ten Results with the Greatest Difference in Ranking Between Marking Methods (Analytical v Pairs) (N=10)

Student ID	Rank analytic/60	Rank Pairs Holistic	Rank Pairs Criterion 1	Rank Pairs Criterion 2	Rank Pairs Criterion 3
ma108	5	43	46.5	49	37
za117	9	33	40	14	39
ma107	12	43	22.5	48	45.5
xa113	19	55	52.5	53	56
ma104	27	51	41.5	51	52
za114	38	17	16	9.5	13.5
wa115	38	20	14.5	24.5	23
za108	44	17	22.5	16.5	20
za107	49	27	26	35	23
la101	52	28	43.5	31.5	31.5

This presents the pairs results of each of the ten work samples according to the specific marking criteria. The correlation between the overall ranking (Pairs Holistic) and Criterion 3 (creative application of appropriate design principles and elements such as alignment, balance, contrast, emphasis, harmony, proportion, proximity, repetition, unity, and white space- focus on brochure and logo) was very strong and significant ($r=0.96$, $p<0.01$). The correlation between the overall ranking (Pairs Holistic) and Criterion 2 (technical proficiency- demonstrable capability and facility with the range of required software -spreadsheet, logo, brochure - focus on features of graphs, logo and layout in brochure) was also very strong and significant ($r=0.88$, $p<0.01$). This supports the view that the pairs holistic assessment places emphasis on the appearance of the product and technical proficiency and marginalises those aspects of the analytical marking scheme such as design, selection of software, file formats, reflection, design standards and conventions leading to the disparity between assessment methods exemplified by these samples.

Comparison between Scores for Portfolio and Examination

This section makes comparisons between the results of marking two different forms of assessment; the Portfolio and the Examination. The marking of these two major components was analysed separately and then compared. The descriptive statistics are shown in Table 4.19. The distribution of scores is presented in Figure 4.8 and Figure 4.9.

Table 4.19

Descriptive statistics for scores from marking for all students.

	N	Minimum%	Maximum%	Mean%	Std. Deviation
Assessor Average Examination	113	20.0	82.0	49.0	14.0
Assessor Average Portfolio	98	6.0	75.0	37.0	17.0
Assessor Average Total (Port & Exam)	96	13.5	74.5	41.2	15.0
Teacher Examination %	74	13.0	90.0	52.5	20.0
Teacher Portfolio %	51	17.1	94.3	59.5	17.5
Teacher Total (Port & Exam) %	26	18.9	93.0	50.0	20.6

Note the high number of portfolios scoring zero marks. These usually corresponded to student work which was missing, either because it was not done or because it was not received. The two zeros in the examination are for students who did not sit. The number of students assessed is explained as follows: a total of 115 students took part in the study and of these 113 sat for the examination; 98 portfolios were received for marking (14 student portfolios from school WA were not received) and 96 students submitted both the examination and the portfolio. Three schools comprising 51 students supplied teacher marks for the portfolio and two schools (26 students) supplied a teacher mark for the examination. Once again it should be noted that the assessment criteria used by the teacher were not made obvious to the researcher.

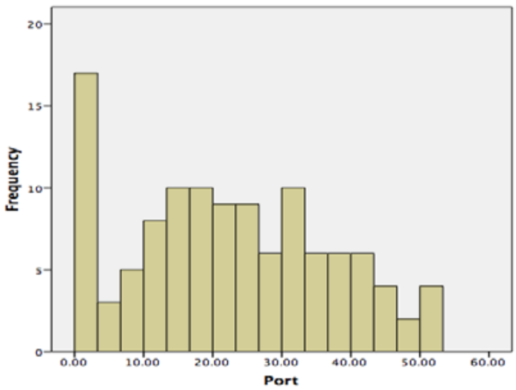


Figure 4.8 Distribution of portfolio marks all cases N=115 (98 submitted portfolios)

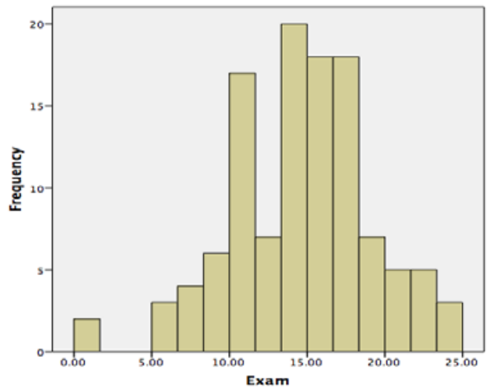


Figure 4.9 Distribution of examination marks all cases N=115 (113 sat examination)

Correlation coefficients between these components are displayed in the Table 4.20 (scores) and Table 4.21 (rankings).

Table 4.20

Correlations for Scores Awarded by Assessors and Teachers for Portfolio (N=51) and Examination (N=26) for All Students for whom Work Samples were Available

	Assessor Average			Teacher		
	Exam	Portfolio	Total (Port & Exam)	Exam%	Portfolio %	Total % (Port & Exam)
Assessor Average Examination	1.00	0.58**	0.75**	0.16	0.40**	0.23
Assessor Average Portfolio		1.00	0.97**	-0.14	0.36**	0.05
Assessor Average Total (Port & Exam)			1.00	-0.07	0.39**	0.10
Teacher Examination %				1.00	0.34	0.62**
Teacher Portfolio %					1.00	0.91**
Teacher Total (Port & Exam) %						1.00

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

Table 4.21

Correlations for Ranking of Scores Awarded by Assessors for Portfolio (N=51) and Examination (N=26) for All Students for whom Work Samples were Available.

Rank of	Assessor Average		
	Exam	Portfolio	Total (Port & Exam)
Assessor Average Examination	1.00	0.57**	0.74**
Assessor Average Portfolio		1.00	0.97**
Assessor Average Total (Port & Exam)			1.00

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

There was a moderate but significant correlation (r=0.58, p<0.01) between the average scores awarded by the external assessors for the examination and for the portfolio. This was in spite of the fact that many submissions of student portfolio work were incomplete. This supports the assertion that candidates' ability may be measured by either a portfolio or an examination though the moderate correlation points to the fact that slightly different qualities are being measured. Correlations between scores awarded by the external assessors and teachers were in general very weak, for example between *Assessors Average Examination* score and *Teacher Examination* score (r=0.16). As mentioned previously, the exact content assessed by the teacher and the method of assessment were not made available to the researcher and so it would be unwise to speculate further upon these results.

Analysis of Results for a Subset Sample Marked Using Comparative Pairs

Further analysis was undertaken of the scores for the 60 students whose examination work was selected for comparative-pairs marking. For this subset, correlations were calculated for the *ranking* of each component (the first three comprising the portfolio; components four and five the examination), as well as the ranking of the mark derived from the pairs marking. These correlations of rankings are shown in the Table 4.22. The correlations between rankings of the portfolio component scores and assessor average examination scores show some strength and significance, particularly for component 1, the portfolio product and component 2 the portfolio design document (r=0.63, p<0.01 and r=0.60, p<0.01 respectively) as shown in the final column of Table 4.22. For component three, the two additional digital artefacts, the correlation, though significant is weak (r=0.38, p<0.01). The correlations between the rank of comparative-pairs scores (component 5) and all portfolio components are weak as shown in the first row of Table 4.22.

Table 4.22

Correlations between Rankings of Average Assessors Scores Analytical Marking (Components C1 C2 C3 Portfolio and C4 C5 Examination) and Comparative-pairs Marking (Component 5 Examination) for the 60 Selected Candidates

Rank of	Assessor Average							
	Pairs Hol.	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Portfolio	Exam
Pairs Holistic (C5)	1.00	0.32*	0.33**	0.17	0.38**	0.72**	0.30*	0.65**
Ass Average Component 1		1.00	0.76**	0.54**	0.66**	0.49**	0.89**	0.63**
Ass Average Component 2			1.00	0.56**	0.65**	0.44**	0.92**	0.60**
Ass Average Component 3				1.00	0.45**	0.27*	0.76**	0.38**
Ass Average Component 4					1.00	0.53**	0.69**	0.82**
Assessor Av Component 5						1.00	0.45**	0.90**
Ass Average Portfolio							1.00	0.62**
Ass Average Exam								1.00

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

Analysis for the Subset of Two Classes with the Most exemplary Implementation

Only two schools (MA and ZA) implemented all aspects of the portfolio (Components 1 to 3) largely in line with the stated requirements and analysis of these 28 students separately is presented below. For these two cases, the range of portfolio marks and scatter plot of examination mark against portfolio mark are shown in Figures 4.10 and 4.11 below. Correlation coefficients between portfolio and examination marks (analytical marking) were much higher for this sub-group of 28 students than for the whole sample of 115 students. For example, the average assessor mark for the examination and average assessor mark for the portfolio are strongly and significantly correlated ($r=0.79$, $p<0.01$). Further, both of these scores correlate highly and significantly with the teacher's assessment of the portfolio, and moderately but significantly with the teacher's overall semester score. The teachers' portfolio and overall semester marks are very highly correlated ($r=0.90$, $p<0.01$) for this sub-sample as shown in Table 4.23. Correlations between scores for the individual components of the portfolio and marks for examination, teacher portfolio mark and teacher semester mark, show much greater strength for this subset of marks.

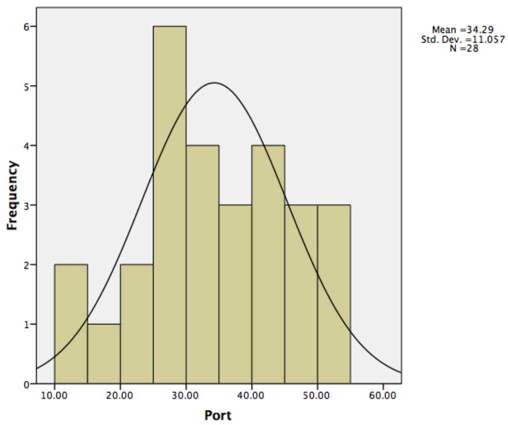


Figure 4.10 Distribution of portfolio marks for schools MA and ZA. (N=28)

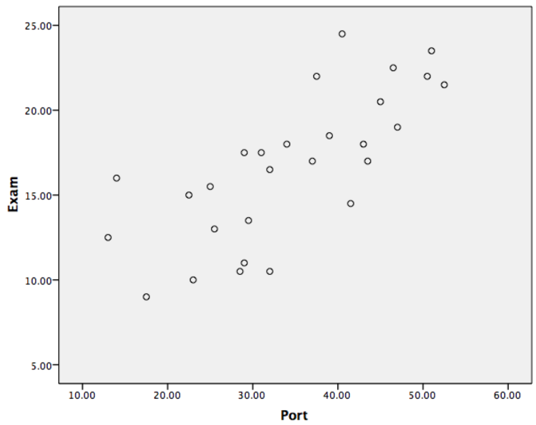


Figure 4.11 Scatter plot of examination marks against portfolio marks for schools MA and ZA

Table 4.23

Correlations Between Marks for the Portfolio (Components 1-3) and the Examination (Components 4 & 5) for Schools MA and ZA. (N=28)

	Assessor Average					Teacher	
	Portfolio	Exam	Comp 1	Comp 2	Comp 3	Portfolio %	Semester %
Ass Average Portfolio	1.00	0.79**	0.74**	0.89**	0.47*	0.70**	0.58**
Ass Average Examination		1.00	0.77**	0.95**	0.80**	0.68**	0.52**
Ass Average Component 1			1.00	0.60**	0.00	0.51**	0.42*
Ass Average Component 2				1.00	0.13	0.69**	0.73**
Ass Average Component 3					1.00	0.17	0.01
Teacher Portfolio						1.00	0.90**
Teacher Semester %							1.00

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

There was a high and significant correlation ($r=0.77$, $p<0.01$) between the Assessors' Average score for the portfolio product (component 1) and for the examination (components 4 & 5). There was a very high and significant correlation ($r=0.95$, $p<0.01$) between the Assessors' Average score for the portfolio design document (component 2) and for the examination (components 4 & 5). There was a high and significant correlation ($r=0.80$, $p<0.01$) between the Assessors' Average score for the portfolio digital artefacts (component 3) and for the examination (components 4 & 5). There were moderate correlations between Assessors' Average scores for the portfolio product (component 1) and Teacher Portfolio mark and Teacher's Semester % ($r=0.52$, $p<0.01$ and $r=0.42$, $p<0.05$ respectively). There were high correlations between Assessors' Average mark for the portfolio product (component 2) and Teacher Portfolio mark and Teacher's Semester % ($r=0.69$, $p<0.01$ and $r=0.73$, $p<0.01$ respectively). All of this supports the view that students' ability was consistently recognised irrespective of the assessor or of the type of work sample (portfolio or examination) with the exception of scores awarded to component 3, the two further digital artefacts. Overall, these results demonstrate that the greater the congruence between what was assessed, externally and by teachers, the greater the reliability of the scores produced.

Rasch Analysis of the Results of Analytical Marking

Rasch analysis of the scores from the analytical marking was conducted using a polytomous model to test the reliability of the judgments for each component of the examination and portfolio as measures of the ability of a student in AIT. Assessor judgements from analytical marking were scored 0, 1 for two ordered categories, 0, 1, 2 for three ordered categories, 0, 1, 2, and 3 for four ordered categories and so on to indicate increasing levels of proficiency in the aspect or skill under test. These responses were then summed to produce a total score for each student. To determine if this total score accurately characterised a particular student's ability and further, if a student with a higher total score than another could be said to be more proficient in the skill set under investigation, a Rasch Uni-dimensional Measurement Model (RUMM) was applied to the marks using the software package *RUMM 2020*. (RummLab, 2011)

Rasch Analysis of Examination Marks

A polytomous Rasch model was applied to the examination scores (both the theory section component 4 and the practical section component 5) using the scores of both assessors to generate a combined score for each student. This resulted in a mean person location of 0.23, fit residual of -0.35 and standard deviation of 1.22. The Separation Index (SI) was 0.85. There were few extreme outliers and the frequency distribution was normal and relatively well spread as represented in Figure 4.12. The correlation coefficients between the location scores and raw marks were high for both assessors ($r=0.90$, $p<0.01$ and $r=0.87$, $p<0.01$) and for the mean of their marks ($r=0.99$, $r<0.01$) as shown in the Table 4.24.

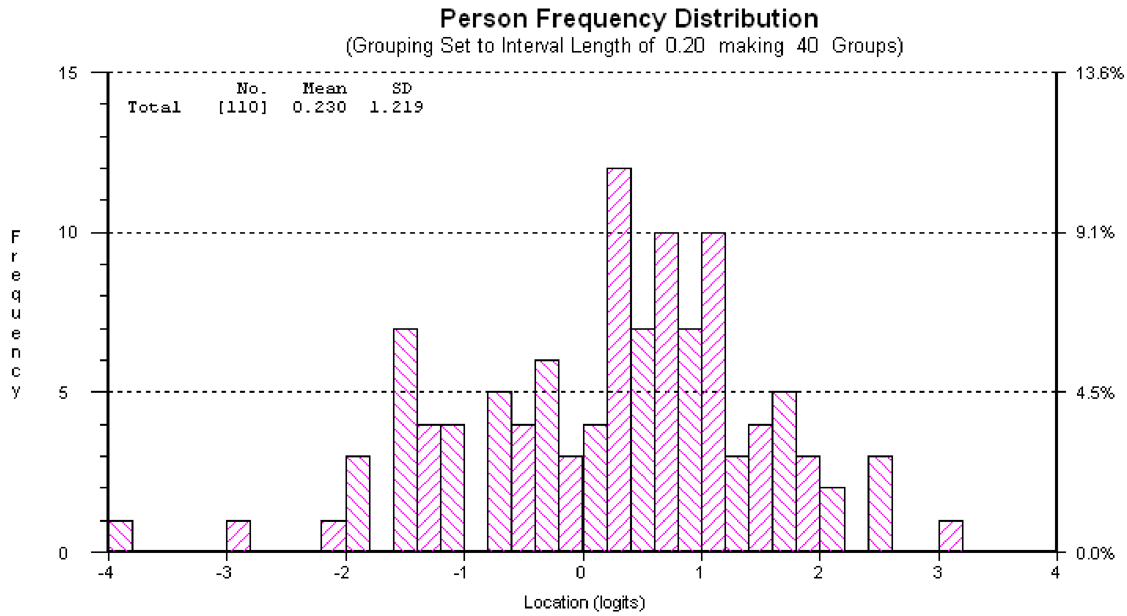


Figure 4.12 Frequency distribution of average examination scores (N=110)

Table 4.24

Correlations Between Location Scores and Raw Marks for Each Assessor and the Average Assessor Mark. (N=110)

	Location	Assessor 1	Assessor 2	Ass Av
Location	1.00	0.90**	0.87**	0.99**
Assessor 1		1.00	0.58**	0.91**
Assessor 2			1.00	0.87**
Assessor Average mark				1.00

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

Figure 4.13 shows the relative location and distribution of all 110 student scores from the analytical marking of the examination. The box plot for each score indicates the standard error.

Rasch Analysis of the Portfolio Scores

A polytomous Rasch model was applied to the portfolio scores using the judgements of both assessors to generate a combined score for each student for each of the three components of the portfolio. For each component, if a student did not submit work they were removed and thus not all the 115 students were included in the analysis. Rasch analysis of the portfolio analytical marks for components 1 to 3 (Product, Process Document and Extra Artefacts analysed separately) gave a reliable set of scores for all three components (SI=0.96, 0.96 and 0.94 respectively). The results are shown in the Figures 4.14 to 4.16 and Tables 4.25 to 4.27.

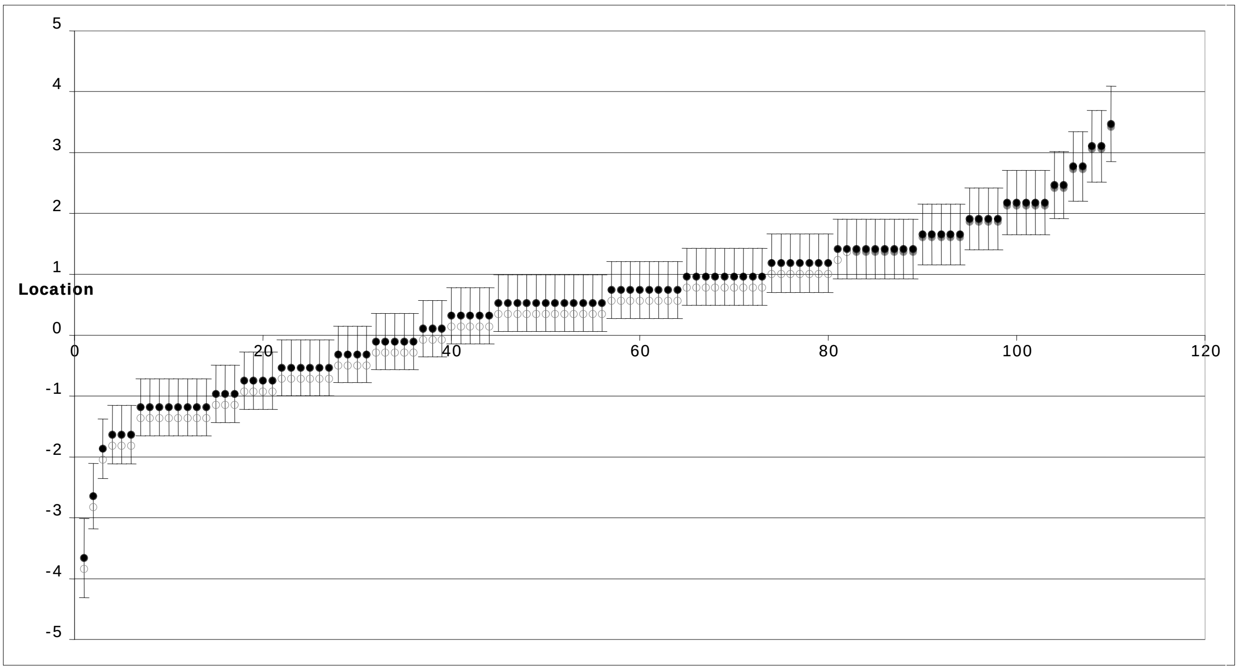


Figure 4.13 Relative location and distribution of student score

Portfolio Product (component 1)

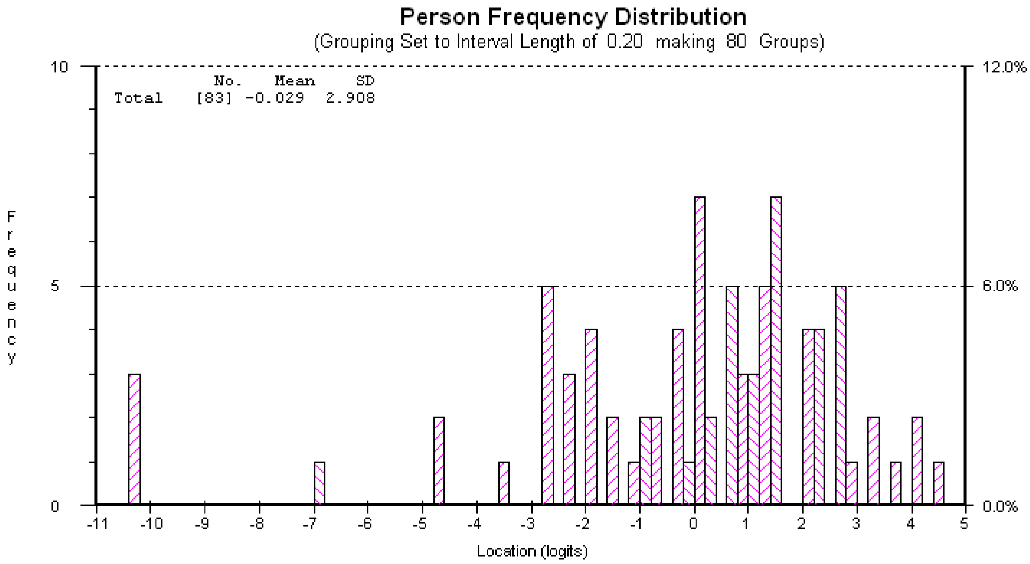


Figure 4.14 Frequency distribution of component 1 scores.

Table 4.25

Correlations between Location Scores and Raw Marks for Each Assessor and the Average Assessor Mark for Portfolio Component 1 the Product (N=83)

	Location	Assessor 1	Assessor 2	Ass Av mark
Location	1.00	0.87**	0.88**	0.96**
Assessor 1		1.00	0.67**	0.92**
Assessor 2			1.00	0.91**
Assessor Average mark				1.00

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

Portfolio Design Process Document (component 2)

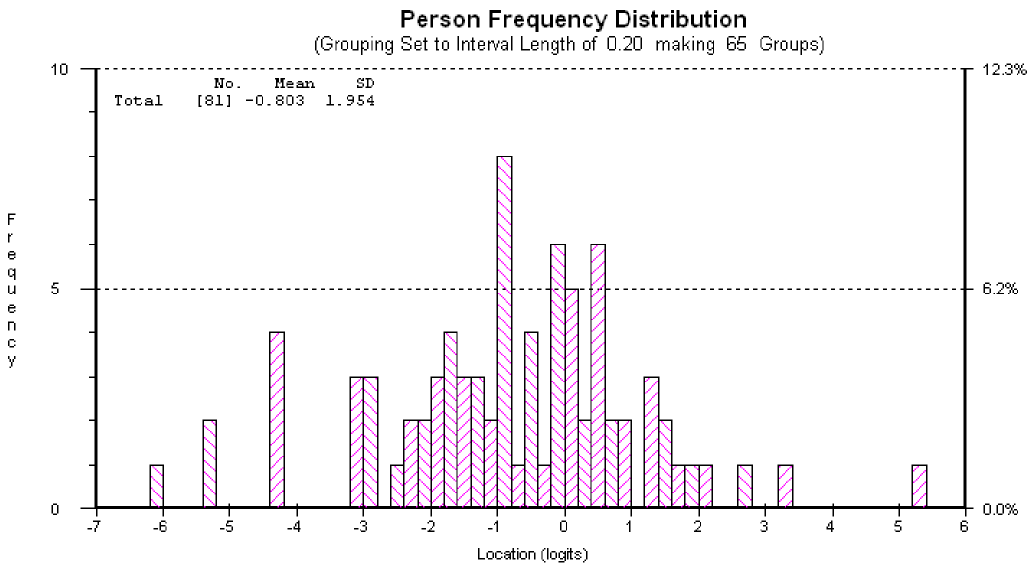


Figure 4.15 Frequency distribution of component 2 scores.

Table 4.26

Correlations between Location Scores and Raw Marks for Each Assessor and the Average Assessor Mark for Portfolio Component 2 the Design Document (N=81).

	Location	Assessor 1	Assessor 2	Ass Av mark
Location	1.00	0.93**	0.94**	0.98**
Assessor 1		1.00	0.82**	0.95**
Assessor 2			1.00	0.96**
Assessor Average mark				1.00

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

Portfolio Two further digital artefacts (component 3)

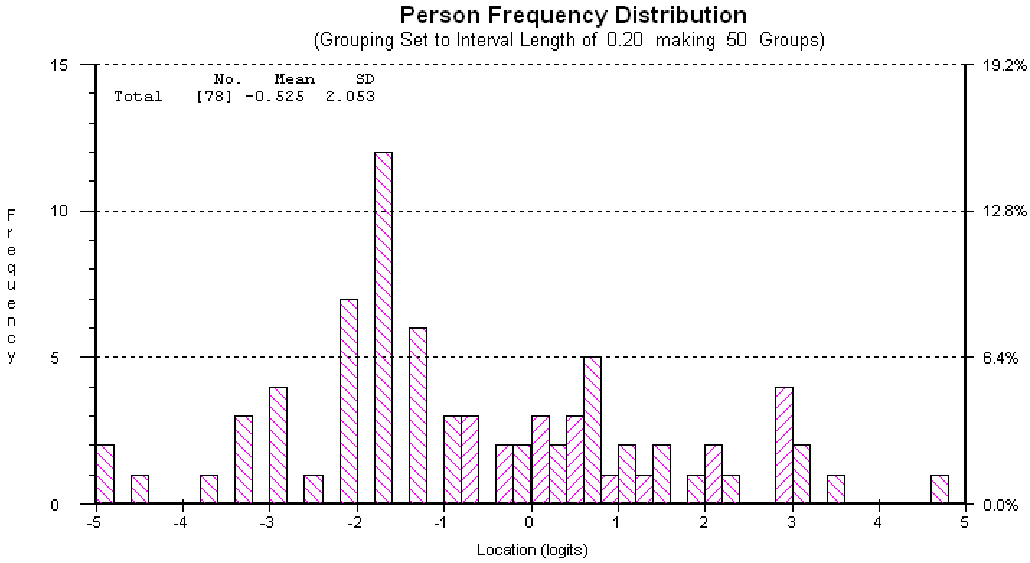


Figure 4.16 Frequency distribution of component 3 scores.

Table 4.27

Correlations between Location Scores and Raw Marks for Each Assessor and the Average Assessor Mark for Portfolio Component 3 the Two Further Digital artefacts (N=78).

	Location	Assessor 1	Assessor 2	Ass Av mark
Location	1.00	0.92**	0.89**	0.99**
Assessor 1		1.00	0.66**	0.91**
Assessor 2			1.00	0.91**
Assessor Average mark				1.00

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

There were a few extreme outliers particularly for the first component, the product. These tended to be students scoring 0 on almost all of the criteria. The frequency distributions tended to be well spread, with high standard deviations and not very ‘normal’ in structure. The very high and significant correlations between *Assessors Average Mark* and *Location*, attest to the strong relationship between the two for all components of the portfolio.

Summary

This chapter described the development and implementation of the assessment tasks and the data collected from observations, surveys, interviews and results of marking. Both the portfolio and examination were developed in close alignment to the course outcomes (see Appendix A) and were perceived to be valid assessment instruments by students and teachers alike. Students' survey and interview data demonstrated strong positive sentiment towards computer based assessment. Teachers' comments during visits and data derived from interviews also pointed to a desire for assessment reform with some satisfaction with the type of assessments trialled here. An analysis of the scores generated by analytical marking supported the assertion that either a portfolio or a computer-based examination may be used to reliably assess student performance in AIT. The strong and significant correlation between the scores generated by analytical marking and those generated by the comparative-pairs marking supports the view that either method is capable of reliably measuring student performance. Rasch analysis of the distribution of the portfolio and examination scores supports the assertion that both instruments were reliable methods of discriminating between student ability in AIT. The next chapter presents a case study for each of the seven schools involved in the project.

CHAPTER FIVE

CASE STUDIES

Introduction

This chapter presents the findings from each of the seven schools, seven teachers and eight classes of Year 11 or 12 students involved in the project. The chapter adopts a multi-case approach (Burns, 1996) with each school forming one of seven separate but parallel experiments from which any common findings may, with some confidence, be generalised to the wider population. Improved internal reliability and validity were promoted by drawing on the combination of qualitative and quantitative data and the perspectives of teacher, student and researcher, involved in the same series of activities in similar environments. The scope of the case studies is briefly introduced in Table 5.1.

Table 5.1

School Type, Participant Numbers and Year Groups

School Code	Sector	Student Numbers	Year group
CA	Government	20	11
LA	Catholic	10	11
MA	Private	12	11
RA	Government	14	12
WA	Private	14	11
XA	Catholic	29	11
ZA	Catholic	16	11

Case Study CA: Public School

The CA case study involved one teacher and a class using an e-commerce website as the product for the portfolio. There were 23 students in the class but only 20 were involved in the study with 3 having withdrawn. The context for the class was Business Information Technology.

Implementation, Technologies and Issues Arising

The researcher either met or communicated using phone and email with the teacher before the students became involved. This was to discuss the process with the teacher and to test the technologies, in particular the use of computers in the school's laboratory with sound recording and a USB flash drive. Implementation differed slightly in this case. Though the portfolio product was an e-commerce website, the teacher legitimately chose an alternative context to the one supplied with the project documentation. The theme was ICT and the Travel Industry and the project, developed by the teacher, was to design a website for the *Fly by Night* airline to include pages on destinations, schedules and an on-line booking form.

The class was conducted in the computer laboratory shown in Figure 5.1.

Figure 5.1 CA computer laboratory.

The students' computers were all less than three years old and well equipped with office and multimedia software. Observation and comment by students suggested that multi-tasking, particularly with large files, often strained system resources on these computers.

Data Collected

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 assessment tasks.

Observations of the Classes

The class was visited on four occasions to observe students completing the assessment task, or to collect qualitative data.

Visit 1: Product Development (15/08/08, 9:05-10:09)

There were 21 students present in the computer laboratory working for about one hour on portfolio component one, the *Product*. There was some initial disruption due to a fire in the administration block on the previous evening and as a result there was no Internet access. The class had commenced development over a week earlier and had just handed in a design document that included their research with a design folio containing detailed hand-drawn storyboards, a timeline, graphic images of logos and answers to evaluation questions. Many of these documents were observed to be very comprehensive, being well written and researched with plenty of detail and up to 20 pages in length. Research was detailed and often included analysis of current airline websites such as *QANTAS* and/or *Virgin Blue*. Observation of students' work revealed a familiarity with the stages of the *technology process*. Students appeared comfortable with developing the documentation and confirmed when questioned that they had been using the *technology process* as a guiding principle in their work since Year 8.

The teacher commenced the lesson by going through the production requirements of the project, reminding the students of deadlines and asking them to start and keep up a daily journal. Only about half the students commenced work diligently and it seemed likely that the fire had distracted the others who took about 30 minutes to settle and really become engaged in the task. A few students were observed developing timelines for production in *MS Word*, graphics such as banners and logos and *Adobe Dreamweaver* to develop web pages. Student said they had used *Dreamweaver* earlier in year and had completed a refresher earlier in the week and had some tutorials available in the form of a written booklet. About 45 minutes into the lesson, all students appeared to be fully engaged with about half working on web pages using *Dreamweaver*.

Visit 2: Development of Process Document (28/10/08, 2:31-3:25)

There were 21 students present in the computer laboratory working for about one hour on portfolio component two, the *Design Process Document*. The lesson began with class members seemingly arriving from all points of this very large campus and perhaps10 minutes had elapsed before all were seated. This was the final period of the day and quite a bit of lethargy was in evidence. The teacher brought up the design process template on the data projector and talked through the requirements for a minute or two. The class then got started with varying degrees of urgency.

A couple of students were immediately on task; others were observed to be tidying up work from other subjects, surfing the internet, looking at the upcoming examination timetable or being generally engaged on their own IT agenda, for example catching up on email, news, other pressing work, prioritising as they saw fit. Eventually after about 15 minutes, most seemed to be working on the *Design Process Document*, however a few (3 or 4) were busy on a brochure for the e-commerce travel website which comprised the theme of their portfolio. Also, it was surprising to see that these students had developed multi-table databases as an activity within the context. The ones observed were very well conceived and correctly linked through key fields. The teacher clearly interpreted this type of activity to be within the scope of the syllabus for AIT 2B although there is no specific reference in the syllabus. As ever, with new courses, there are different interpretations of the course outline. Students populated the design document using cut and paste from the work they had done along the way. They added in their investigation notes of e-commerce sites and their design ideas which were typically scanned sketches together with notes on their production process and evaluation of the final product.

The teacher commented that students were finding difficulty in selecting what to add to each section of the *Process Document*. Students seemed understandably reluctant to re-write what they had already done along the way. The teacher and researcher discussed the arrangements for research components four and five, the examination. The examination paper itself had been modified after the first implementation at RA with the order being practical (component 5) before reflective/ theory questions (component 4) making possible the completion of the questionnaire and student focus group on same day. The exam paper had already been printed and scheduled for 23rd November in a different room to the current one with later model computers. The class were quite impressed with the headset, which the teacher demonstrated, and students were informed that subject to returning signed consent forms they would be allowed to keep the one issued to them for the examination. This prompted a flurry of requests for permission forms. No issues with audio recording were anticipated by the teacher.

Visit 3: Examination (25/11/08 8:50-1:00)

Twenty-three students sat the 3-hour, two part examination in a computer laboratory which was different to the normal room for the class. The examination began with a sound test followed by 10 minutes reading time during which students were allowed to browse the files on the USB flash memory. Students were provided with a temporary and individual logon (e.g. CA###, Exam01) that gave them access to their Portfolio Product (a website) only and allowed internet connectivity to be restricted for the duration of the examination. All USBs worked and were correctly imaged but the sound recording test using *MS Sound Recorder* apparently did not for a number of students so it was decided to abandon the audio recording with students simply extending their reflective notes. At the time it did appear that no students had been able to make an audio recording; however, later a number of students said that they could. The teacher said that he had tested the sound recording before and it had worked so it was most likely that the problems were with the students themselves.

Students began on time at 9am for the 3-hour examination and were instructed to work on Task 1 only (Planning) for the first 15 minutes using either the templates provided on the USB drive or on the paper copies. Three students were observed to open other applications during this planning time and were told to exit these. A number of questions were fielded during the first minutes, for example, "Are we allowed to use the photos on the USB in planning?" (Yes), "Can I do one design on Computer and one on paper?" (Yes), "Can I use photos off the Internet?"(No). Most students did some of their planning with the computer. It should be noted that the teacher had changed mark allocation on the paper copies of the examination and that these were different to those on the copy of the paper supplied on the USB and he informed the class of this fact during the opening minutes. As the examination proceeded students were observed working using the following application software: *Paint*, *CorelDraw* and *PaintShop Pro*. Minimal planning was observed to be done on the *MS Word* and *PowerPoint* templates supplied. The brochure (the end product of the practical activity) was observed to be exclusively done with a *MS Publisher* template.

The first technical issues occurred after about half an hour. Student CA107's computer stopped responding and had to log off and be restarted and then had difficulty re-opening a JPG file. Student CA105's *Publisher* application stopped responding and observation revealed that, entirely understandably, a lot of windows were open suggesting that these machines were struggling to cope with multiple applications. Student computers were very close together and there were 23 in quite a small room. With no privacy screens in place, some *lateral diffusion* of ideas was evident with neighbouring candidates creating similar shapes and adopting similar colours schemes for component parts of the product. A few problems creating graphs were also evident. The problem with audio recording resurfaced when a student asked "if my sound *is* working what do I do?" A general announcement was made and a show of hands indicated that half the class did in fact have audio recording functioning correctly.

Transition to component 4, the reflective questions, was achieved smoothly after 2 hours working time on component 5, though some candidates were still finishing off and had to be instructed to save and move on. Almost immediately, a new set of questions had to be fielded. Principally these were requests for explanation of wording and were of the type *what does this mean?* This suggested that the language employed in the reflective questions might have been too advanced for these students and that some examples might have been of assistance in clarification.

Visit 4: Student Survey and Forum (25/11/08 1:00-2:00)

On completion of the examination, students were presented with, and completed, a questionnaire. Two student forums (consisting of 7 students and 6 students) were convened by invitation of the researcher and on a voluntary basis. Each group was presented with the same set of structured interview questions with follow up questions differing according to responses.

Survey of Students

The survey was delivered immediately after completion of the performance examination. The minimum, maximum, mean and population mean (mean for all 115 students across the seven cases) were calculated for 20 students' closed response items using SPSS (refer to Appendix K).

Items Concerning the Portfolio

Though strongly positive, students in this group were slightly less enthusiastic about using computers for the portfolio when compared to the population as a whole. Means for items concerning responses to the portfolio ranged from 1.9 to 2.3 on a scale of 1 (strongly agree) to 4 (strongly disagree) as compared to means of 1.6 to 2.1 for the population. Students were slightly less confident about the amount of time they would require to become familiar with computer-based portfolio assessment with a mean of 2.4 on a scale of 1 (lots of time) to 4 (no time) as compared to 2.7 for the population. Students either agreed or strongly agreed with the assertion that the computer was easy to use for developing and presenting portfolio ideas and creating and reflecting on portfolio products; means ranged from 1.9 to 2.3 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students were also unfailingly positive about the assistance provided by the structure or *steps* in the portfolio and asserted that they were able to adequately demonstrate what they could do. Means for these responses were 2.1 and 2.3 compared to population means of 1.9 to 2.0 indicating that this group was slightly less positive than the sample as a whole.

Items Concerning the Examination

Student responses suggested that they had not done examinations on computer before (15 students had no experience and 6 only minor experience) but 12 students indicated it would take only a little or no time to become accustomed to doing so. Again, in line with the portfolio product, students in this sample were slightly less positive than average about the time required to become used to computer-based examinations with a mean of 2.6 compared to 2.8 on a scale of 1 (lots of time) to 4 (no time). Students were firmly in favour of using computers for all aspects of the examination (in preference to a paper based test) as indicated by strong positive sentiment in response to items pertaining to presenting ideas, creating logos, graphs and brochures, reflecting on design ideas and designing products, with means ranging between 1.6 and 2.5 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students were also positive about the assistance provided by the explanatory structure or *steps* in the examination (mean=2.2) and asserted that they were able to adequately demonstrate what they could do in the examination (mean=2.1). Again this group was slightly less positive than the population as a whole where means for these items were both 1.9.

Items Concerning Computer Use

Students indicated widespread ownership and usage of digital devices outside school with 100% broadband internet connectivity reported. These students were regular and extensive users of new technologies with almost all using a computer at home on a daily basis, sometimes for communication or educational purposes, as well as a full range of home entertainment systems. *Attitude* to computers was definitely positive with widespread facility of use reported. Students' self-assessment of proficiency with application software on a scale of 1 (low) to 4 (high) was also elevated. Means ranged between 3.1 and 3.8 for with the exception of databases, spreadsheets and digital video editing where means ranged from 2.5 to 2.8.

Open-ended Items

There were four open-ended questions on the student questionnaire. Responses to the open-response items were tabulated to assist in drawing out themes. There were various responses and a sample of these is shown in Table 5.2.

Table 5.2

CA Summary of Students' Responses to Four Open Ended Questions on Survey Questionnaire

The two best things …	The two worst things …
Portfolio	Portfolio
Easy and convenient to work, edit, store and design (16 responses used the words easy or easier) e.g. "easy to create websites" "It was easy to fix errors". Learning new skills. (6 responses). E.g. "learning and array of programs". "The chance to learn something new'.	Difficulties, inconveniences and fears of malfunction were cited. 14 students used the word hard or difficult in relation to some aspect of the portfolio. E.g. "hard to study when there are no physical notes". "Can be difficult". 9 students alluded to technical fears e.g. "Computer can crash and lose files". "It is at risk of being corrupted or deleted"
Examination	Examination
Easy (12 responses included the words easy or easier) e.g. "I could easily edit any mistakes I made" Appropriate (4 responses alluded to the suitability of the exam) e.g. "I could actually show the examiners" and "fitting to the subject" Convenient (3 students) e.g. "I didn't have to bring anything to the exam" and "typing is more convenient"	Reliability (10 students mentioned fear of technical difficulty associated with hardware and software) e.g. "Possible technological malfunctioning" Structure of the Exam (3 students criticised the nature of the exam) e.g. "very linear - in normal exams you can do the sections in an order that suits you" Disturbance (3 students mentioned distractions) e.g. "disturbance of instructors telling us what to do"

Students were asked to list the two best things about doing the portfolio and examination by computer and to list the two worst things about doing the portfolio and examination by computer. 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 worst things were a concern that the computer could crash and their work might be lost.

Questionnaire Scales

Some of the results of an analysis of the seven scales derived from combining selected items from the questionnaire are shown in Figure 5.2 and Table 5.3.

Table 5.3

CA-Descriptions and Descriptive Statistics for Scales Based on Items from Student Questionnaire

	N	Min	Max	Mean	SD	Description
eAssess	20	1.36	3.55	2.95	0.48	Perceived efficacy of computer use for the examination Potential range between 1 and 4
eAssessP	20	1.18	3.36	2.86	0.48	Perceived efficacy of computer use for the portfolio. Potential range between 1 and 4.
Apply	20	1.60	3.00	2.20	0.45	Application of computer to various uses. Potential range between 1 and 3.
Attitude	20	1.40	3.00	2.53	0.34	Attitude towards using computers. Potential range between 1 and 3.
Confidence	20	1.50	3.00	2.64	0.34	Confidence in using computers. Potential range between 1 and 3.
Skills	20	1.91	4.00	3.14	0.55	Self-assessment of ICT skills. Potential range between 1 and 4.
SCUse	20	0.0	192	71	412	Estimate of time in minutes per day spent using computers at school.

An explanation of the scales is set out in Table 3.4 in Chapter Three. Results indicated that the examination and portfolio were both perceived as appropriate and relatively easy to complete with means *eAssess* and *eAssessP* approximately one standard deviation above the mid-point. *Attitude* and *confidence* scales were even more positively skewed with means approaching two standard deviations above the midpoint. Students had a high, self-assessed level of ICT skills across a range of applications (mean=3.14, midpoint=2.5, SD=0.55). On average these students indicated using ICT for a little over 1 hour per day at school.

Student Forum

Two student groups were interviewed immediately after the examination with discussion focussed on the examination and not the portfolio The first group had finished early and the second agreed to stay behind immediately after the examination.

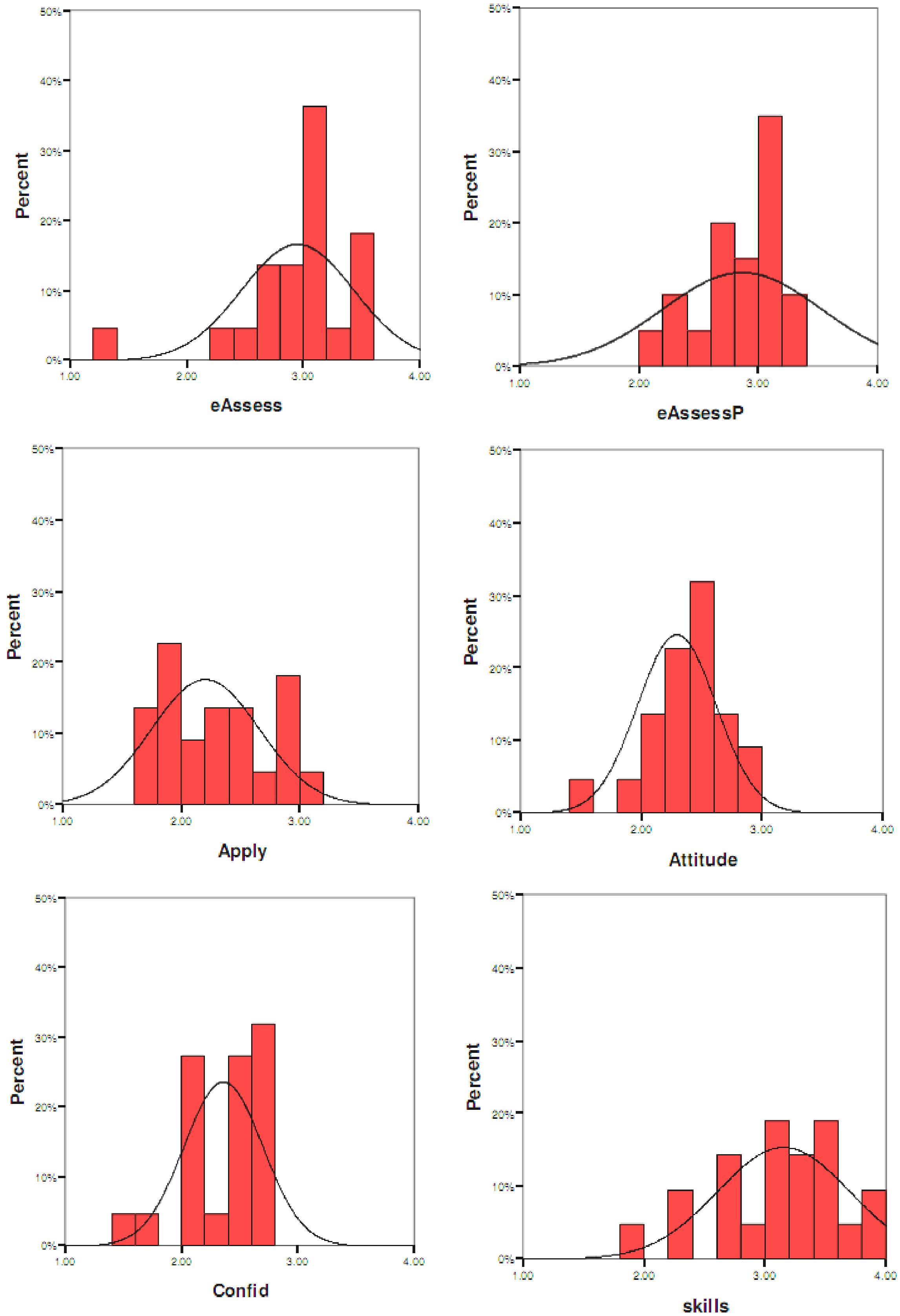


Figure 5.2 CA graphs for the distribution of scores for scales on the student questionnaire.

The first discussion ran for 4m 59s and the second for 6m 19s.

Group 1

What did you think of the task(s) you were asked to do?

"quite long but quite simple to do" "weren't difficult but just the time" "they were good tasks but the time was too short." When prompted, students suggested an additional half an hour would be required.

Did the computers help?

It was generally agreed that given the time available the work was of acceptable quality. "It might have been a bit rushed but besides that I thought I actually did quite well"

Did the computers help? A chorus of "yes" "definitely easier"

How much different was this to how it used to be done?

"This is a lot better" "A lot simpler" "because you can type there's no messy papers and if you make a mistake you can like backspace- you don't have to worry about crossing out" "you just go bang-I want it there and it's done". All agreed they were able to produce a lot more in the given time.

What, if anything, would you like changed in future?

"More time". All concurred with this sentiment. "Or just less to do".

Were there any technical problems with doing the activities?

There were major issues here with sound recording. "That's my only downside using the computers- if something screws up you don't have anything else" "I had to shut down programs three or four times".

Were there any other problems with the activities?

"The wording in the second part…a bit confusing. I had to guess at what it meant" Students would like simpler, clearer instructions and felt the language used was beyond them.

Any other thoughts … or suggestions for developing the use of digital forms of assessment?

“Having newer computers …it took a while to load applications” and this was a constant worry for students.

Group 2

What did you think of the task(s) you were asked to do?
“Too loaded, too many questions I didn’t finish” “Easy…wasn’t very challenging, just time consuming”. “I think the practice exam [company annual report] was harder” “It was good because it was easy” [laughter].

It was generally agreed that quality work could be done. A short discussion over the timing ensued with some students saying there was enough or too much time and others too little.

Did the computers help?
A chorus of “Yes. Much easier”.

How much different was this to how it used to be done?
“This is a lot better”. When prompted all agreed this was a fairer method of assessment and one that they enjoyed doing.

What, if anything, would you like changed in future?
“More time and less questions”. “Get rid of the audio section- even though it didn’t work” All wanted to be allowed freedom to choose their timing. “Give us example time but let us choose”. “People know how they work best”.

Were there any technical problems with doing the activities?
“My page closed before I’d saved”. There was considerable lagging with some applications here and a few crashes. Graphing was not understood by some of this group. “Graphs...I didn’t know how” [laughter].

Any other thoughts or suggestions for developing the use of digital forms of assessment?
“Less reflective questions.” There was agreement that much of this had been done already over the year.

In summary, Both groups were positive about computer-based assessment with the only caveats being the potential for technical difficulties and the time pressure perceived to be the result of the requirement to complete too many activities. The sentiment of both groups was firmly in favour of a computer examination as a fairer and more authentic assessment method. Students were critical about the rigidity of the timings for each section and the subsections of the practical, and of the quantity of the reflective questions.

Email 'Interview' with Teacher

What did you think of the assessment tasks overall?

> The assessment tasks were really good. There was a problem with the compiling of what my students did into the written Design Process document. I would have preferred to submit my document requirements instead.

What did you think of the structure of the activities?
"I liked the structure. The students were able to demonstrate many aspects of the syllabus".

What were the students' reactions to the activities?
"They were happy enough with the practical but I got a feeling they weren't all that keen about a practical exam".

What do you think of its potential?

> After marking the TEE AIT exam, a practical exam for the production component couldn't come quickly enough. If the external marking of a portfolio does away with the moderation process, I'm all for it.

What did you think of the quality of work produced by your students for these tasks?
"I felt I got better results from the students this year as opposed to last year because of the wording of the tasks supplied by you guys".

Were you surprised by the performance/attitude of any students? . "After 36 years of teaching, nothing surprises me"

What was the general feedback from students?
“They were happy enough with the tasks. They dragged their feet with the process document. I’m not sure about the exam”.

Were there any technical problems with implementing the activities?

> The headphones didn’t work during the exam. I checked the computers by logging into an exam login and tested the sound with my headphones. Everything worked. On the day of the exam, most students could not record with the supplied headphones. Please forward a pair to me so I can test them.

Were there any other problems with implementing the activities?
“I could not read the photocopy of one of the student’s design sheet. It had to be rescanned and sent”.

Any other thoughts or suggestions for developing the use of digital forms of assessment? “Maintain the rage!”

In summary, the teacher was positive about the nature and structure of the tasks for both the portfolio and examination perceiving both as valid and authentic and satisfied that these allowed students a fair opportunity to demonstrate their ability. A couple of technical issues were noted, principally the failure of some students to record audio and the illegibility of photocopies of student design sheets for the examination.

Results of Marking

Table 5.4 shows the scores awarded by the two external assessors and by the teacher. The two external assessors marked five pieces of work for each student comprising the design document for the product, the product itself, two further digital artefacts, a theory section and a practical section of a three-hour examination. The marks for these were totalled. No allowance was made for missing work. Teachers were requested to mark the examination (both theory and practical sections) and award a semester mark for the student. In this instance no examination marks were supplied. The *rank-assessor average* is the position out of the whole sample of 115 students spread over seven schools. The *rank-teacher’s semester (Tch Sem)* is the rank only within the class. The mean of the average of the external assessors' marks (32.2%) was not statistically significantly different to the mean for all students involved in the study (36.9% with SD of 17.2%). There was only one student in this class ranked in the top 10% of students (*Rank Ass Ave*). The mean ranking of the external assessors was 67.2, above the mid-point of 58, indicating that overall the scores were low.

Table 5.4

Results for Case CA from Marking Portfolio and Examination (N=20).

St ID	Assessors marking (Total %)			Assessors		Teacher (%)	Rank	
	Ass1	Ass2	Ave	Pfolio (70)	Exam (30)	Sem	Ass Ave*	Tch Sem
ca101	37	30	33.5	23.0	10.5	70	63.0	3.0
ca102	33	13	23.0	12.5	10.5	38	86.5	12.0
ca103	21	31	26.0	26.0	b	13	81.0	19.0
ca105	61	56	58.5	41.0	17.5	43	16.5	10.0
ca106	16	12	14.0	a	14.0	46	105.0	8.0
ca107	32	30	31.0	24.0	7.0	56	70.5	4.0
ca108	30	27	28.5	17.5	11.0	55	75.5	5.0
ca109	32	31	31.5	15.0	16.5	33	68.5	14.0
ca110	23	18	20.5	9.5	11.0	15	89.0	18.0
ca111	51	43	47.0	30.0	17.0	82	32.0	1.0
ca112	24	15	19.5	10.0	9.5	26	91.5	17.0
ca114	30	30	30.0	15.5	14.5	50	72.5	7.0
ca115	22	22	22.0	15.0	7.0	9	88.0	20.0
ca116	31	38	34.5	21.0	13.5	54	62.0	6.0
ca117	62	72	67.0	52.5	14.5	76	5.0	2.0
ca119	23	23	23.0	12.0	11.0	27	86.5	16.0
ca120	33	32	32.5	19.0	13.5	37	66.5	13.0
ca121	42	52	47.0	30.5	16.5	43	32.0	10.0
ca122	39	36	37.5	31.5	6.0	43	55.5	10.0
ca123	18	16	17.0	6.5	10.5	31	96.0	15.0
Mean	33.0	31.4	32.2	21.7	12.2	42.4	67.2	
SD	12.6	14.9	13.4	11.6	3.5	19.4	26.4	

* *Ranking of external assessors is for all 115 students involved in the study.*

a. *Portfolio not submitted.* b. *Student absent for examination.*

A correlation analysis was done on the scores and on the rankings generated from the marking, the results of which are shown in Table 5.5 and Table 5.6. Correlation between the external markers was strong and significant with an inter-rater reliability correlation coefficient of 0.88 ($p<0.01$) on the scores of students.

Table 5.5

CA Correlation Coefficients from the Marking of Students' Work (N=20).

	Assessor 1	Assessor 2	Average	Teacher Semester %
Assessor 1	1.00	0.88**	0.97**	0.64**
Assessor 2		1.00	0.98**	0.53*
Average			1.00	0.60**
Teacher Semester %				1.00

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

Table 5.6

CA Correlation Coefficients from the Ranking of Student Work (N=20).

Rank of	Assessor 1	Assessor 2	Average	Teacher Semester %
Assessor 1	1.00	0.86**	0.96**	0.65**
Assessor 2		1.00	0.96**	0.52*
Average			1.00	0.63**
Teacher Semester %				1.00

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

When compared with the average of the two assessors, the correlation of the teacher's mark and rank for the semester were moderate but significant ($r=0.60$, $p<0.01$ and $r=0.63$, $p<0.01$ respectively). No separate examination results were supplied and the teacher's semester mark was made up of several components only some of which were the external assessment tasks.

Conclusions

Conclusions pertaining to this case study are discussed separately for the portfolio and the examination.

Portfolio

The student portfolios and artefacts were problematic with many submissions being incomplete. This meant that high assessor scores were rare. These components may or may not have been available to the teacher and possibly account for the disparity between assessor score and teacher's semester mark. There was a high level of agreement between the two external assessors though some initial anomalies presented due to inability to display content when marking remotely. The students and the teacher were familiar with this form of assessment, with understanding, skills and knowledge being demonstrated through performance on activities that had an obvious connection to the real world. The portfolio tasks mirrored typical class activity and were easily manageable.

Examination

Capture of the digital content of the examination presented a few difficulties with file types; however with the exception of the audio recording of the students' reflection, the computer supported production examination was implemented with no significant technical difficulties with all students able to complete the requirements of the exam in the time permitted. Malfunctions for two students were quickly rectified by the teacher. However, students felt they needed more time for the practical component. The students responded well to the style of examination and appeared to enjoy the practical component. Many students didn't seem to appreciate the difference between a logo and a banner or poster. Numerical data were handled well with a spreadsheet. Although agreeing that the theory section was made easier by being done on a computer, many students found the reflective questions difficult to understand and repetitive in nature. Many appeared to misunderstand the subtleties of the questions and to which part of the technology process they referred as evidenced by the many answers that were off topic.

Case Study LA: Private School

The LA case study involved one teacher and a class using a website as the product for the portfolio. There were 22 students in the class, which was a mixed group of stage 1 and stage 2 students with 10 stage 2 students being involved in the study. The context for the class was Business Information Technology. The teacher followed the design brief supplied, modifying the context for the portfolio product to an *Olympic Games candidate website*, but as an addition to other coursework and not as the coursework itself. This extra work may have put time pressure on students perhaps preventing them from delivering completed solutions as evidenced by the incomplete submissions from some students. The performance examination was also undertaken as an additional task and not counted towards the student's final semester mark.

Implementation, Technologies and Issues Arising

The researcher either met or communicated using phone and email with the teacher before the students became involved. This was to discuss the process with the teacher and to test the technologies, in particular the use of computers in the school's laboratory with sound recording and a USB flash-drive. The class was conducted in a computer laboratory pictured in Figure 5.3. The students' computers were all less than three years old and well equipped with up to date office and multimedia software. Observation and discussion with the teacher and students indicated that the hardware was well matched to the demands of the software on these computers.

Figure 5.3 LA computer laboratory

Data Collected

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 assessment task.

Observations of the Classes

Members of the research team visited the class on four occasions to observe students completing the assessment tasks, or to collect qualitative data.

Visit 1: Product Development (18/09/08, 2:15-3:10)

There were 10 students present in the computer laboratory working for about one hour.

The room was shared with 12 AIT Stage 1 students working on a different task. The class was addressed to explain the nature of the project and the consent forms. Work had already begun the previous week. Students worked independently but could discuss with each other. Some were completing research, creating concept maps (using online tool), and design documents (e.g. file structure, screen designs, navigation) though most were doing screen designs using paper and pencil. It was suggested to the students and teacher that these could be simply scanned or photographed for the process document.

Two students were working on first screen of a website using *Dreamweaver*. One showed another how to set up frames. Research appeared to centre on comparing 2016 applicant sites and past Olympic websites. Three students worked on timeline spreadsheets for the task. These appeared very detailed, probably too much so and were taking too long. One boy was redoing a drawn design in *MS Word*. It was suggested by the teacher that he just scan it to save time.

An e-mail was received from teacher (12/9/08), “Students would like to do designs e.g. thumbnails, master design, site structure at home for homework – is this acceptable or does all work have to be done in class? What about research of sites that they cannot access due to blocks at school – can these be researched at home?” The teacher was assured that the use of home time by students was permissible.

Visit 2: Development of Design Process Document (31/10/08, 8:45-9:30)

There were 10 students present in the computer laboratory working for about one hour. The room was shared with 12 AIT stage 1 students working on a different task. Students were all engaged working on the Design Process Document having, mostly, completed the website *Perth 2016 Olympic bid*. Students are happy to show their websites in action. These were done in Dreamweaver and students had added original logos, rollovers, and effects to give them some interactivity and visual appeal. The class seems well on task with everyone working away. Mostly they are using the *MS Word* template and populating this with scans of designs done on paper and screenshots of their development process with some explanations of the research they had done, for example looking at previous Olympic sites.

The details of the project were discussed with the teacher and the requirements of the portfolio explained once again, these being the product, a design document plus two other ‘digital artefacts’. Details of the examination were also covered and all appeared to be well understood. The teacher offered to burn the student portfolios to DVD. The examination was scheduled for 21st November at 12:50 and was to be an extra activity to accommodate the requirements of the project. The teacher stated that she wanted to give the students preparation for the following year when it was expected that they would sit an external written paper.

As a prelude to the practical examination, a sound test was undertaken with two students. Both went immediately to *Adobe Soundbooth* and appeared to be quite familiar with this application. Both also knew about *Windows Sound Recorder*. Before leaving, the class was again addressed about the research, the examination and the focus group which many seemed keen to be involved in. This was a very pleasing group to be with and everything seemed to be running smoothly. The practice examination task and final examination, organised with practical preceding theory, were to be sent out to the teacher. The questionnaires and focus group could be completed on the day of the examination on 21st November.

Visit 3: Examination (21/11/08 12:50 - 4:30)

Nine students sat the examination in the computer laboratory. A second examination involving year 10 students was also taking place in the room. The room was spacious enough to allow separation of the students by one workstation. The sound test presented issues for some students and for some, connections to the sound card had to be made from the rear of the system unit. One student was restricted to a particular machine which forced a re-arrangement of the seating. One student was absent.

After a brief explanation of the examination structure reading time commenced. The examination started with the 15 minute design section. All students designed on paper and about half continued to design after the mandatory 15 minutes. Students LA103, LA104 LA108, LA109 and LA110 used *Photoshop* for the logo. Many appeared to take one of the photographic images supplied and add some text to produce a logo. Students LA103, LA104 and LA107 drew free hand in *Illustrator*. Students LA109 and LA110 selected a *Publisher* template for the brochure. Students LA106, LA107 and LA109 used *MS Word* to create a table for room rates from scratch instead of converting from the table supplied using the text to table feature.

After about 50 minutes, some disturbance from questions from the other group was apparent. It was not ideal having a second exam going on in same room. About 80 minutes in, student LA109 started preparing the audio reflection. This audio section constrains the exam, coming in the middle. Perhaps it would be better right at the end after all. The Audio reflection commenced after 115 minutes with some shyness and giggles then silence. A problem occurred with student LA104. Sound was not recording so the student was relocated to a spare machine. Student LA108 forgot to press the record button; all in all a bit chaotic. After 2 hours, students commenced the reflective questions of section B. All students wrote copiously. It was difficult to imagine that they would have produced as much with a pencil. It wouldn't be physically possible to write at this speed. All students used touch typing to varying degrees of proficiency. It's a pity the questionnaire wasn't also computer-based as it will be hard to get them to focus on that after a long exam. With 20 minutes remaining, at least half of the students had finished the examination and started on the questionnaire. Another typographic error was noticed on the exam submission list.

Visit 4: Survey and Forum (21/11/08 4:00 - 4:30)

On completion of the examination, students were presented with and completed a questionnaire. A group of four students agreed to take part in a student forum. They were presented with the same set of questions as other cases but follow up questions differed depending on responses.

Survey of Students

Ten students took part in a survey by completing a questionnaire consisting of 70 closed response items and four open-response items. The survey was delivered immediately on completion of the performance examination. The minimum, maximum, mean and population mean (mean for all 115 students) were calculated for each closed response item using SPSS (refer to Appendix K).

Items Concerning the Portfolio

Students in this group were strongly positive about using computers for the portfolio with responses to in line with the population as a whole. Survey items with positive statements about the portfolio were strongly supported with means ranging from 1.4 to 2.1 on a scale of 1 (strongly agree) to 4 (strongly disagree) as compared to means of 1.6 to 2.1 for the population as a whole. Students were slightly more confident about the amount of time they would require to become used to computer-based portfolio assessment with a mean of 2.9 on a scale of 1 (lots of time) to 4 (no time) as compared to a population mean of 2.7. Students either agreed or strongly agreed with the assertion that the computer was easy to use for developing and presenting portfolio ideas and creating and reflecting on portfolio products; means ranged from 1.7 to 2.1 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students were also very positive about the assistance provided by the structure or *steps* in the portfolio and asserted that they were able to adequately demonstrate what they could do. Means for these responses were both 1.8 compared to population means of 1.9 and 2.0 indicating that this group was slightly more positive than the sample as a whole.

Items Concerning the Examination

Students' responses indicated a range of prior experience with examinations on computer with two students indicating some experience, five little and three no experience. There was similar variation in students' estimate of the time required to become accustomed to computer-based examinations with three students indicating some time would be required and seven little or no time. A mean of 3.1 compared to population mean of 2.8 on a scale of 1 (lots of time) to 4 (no time) suggests that students in this sample were slightly more confident than average. Students were firmly in favour of using computers for all aspects of the examination (in preference to a paper based test) as indicated by strong positive sentiment in response to items pertaining to presenting ideas, creating logos, graphs and brochures, reflecting on design ideas and designing products, with means ranging between 1.8 and 2.1 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students were also positive about the assistance provided by the explanatory structure or *steps* in the examination (mean=2.0) and asserted that they were able to adequately demonstrate what they could do in the examination (mean=1.7). Again this group was slightly more positive than the population as a whole where means for these items were both 1.9.

Items Concerning Computer Use

There was widespread ownership and usage of digital devices outside school with 9 out of 10 reporting home broadband internet connectivity as well as a full range of home communication and entertainment systems; all had a game console and a mobile phone. These students were regular and extensive users of new technologies with a highly positive attitude to computers; all students reported using a computer at home for educational purposes and all responded yes to the statement *I feel confident about using computers*. Students' self-assessment of proficiency with application software on a scale of 1 (low) to 4 (high) was also elevated. Means ranged between 3.3 and 4.0 for a range of applications including spreadsheets, databases and digital video editing, placing this group above the population where corresponding means for these applications ranged from 2.7 to 3.0.

Open-ended Items

There were a variety of responses and a sample of these is shown in Table 5.7.

Table 5.7

LA Summary of Students' Responses to Four Open Ended Questions on Survey Questionnaire

The two best things …	The two worst things …
Portfolio No single theme was evident. Students alluded to 'freedom' e.g. "I was free to use my own ideas" and "It was good to develop my ideas", 'easiness' e.g. "easy to complete" and "I could get it done much faster" and novelty, e.g. ""different, creative" and "New".	Portfolio Six students made reference to difficulties in terms of time and workload. For example, "extra work" and "long reports" and "too much written work".
Examination (8 responses made reference to a preference over handwriting) e.g. "easier to type than write - takes a lot less time". Creative (3 students alluded to the creativity) e.g. "It is possible to create things instead of just designing" and "Fun and creative and familiar". Effective (3 students alluded to the effectiveness the computer) e.g. "I can personally get ideas down on a computer" and "A lot more effective"	Examination Technical Difficulties (6 responses alluded to real or potential problems with hardware) e.g. "concern that hardware may fail" Voice recording (2 students commented on the audio recording) e.g. "Everyone can hear you speak for the audio section"

There were four open-ended questions on the student questionnaire. Students were asked to list the two best things about doing the portfolio and examination by computer and to list the two worst things about doing the portfolio and examination by computer. Responses to the open-response items were tabulated to assist in drawing out themes. Generally students considered that using computers in an examination made for a fairer form of assessment, allowed demonstration of a full range of skills and allowed them to write more fully. The main worst things were concerns about hardware failure and losing work through a software crash.

Questionnaire Scales

Some of the results of an analysis of the seven scales derived from combining selected items from the questionnaire are shown in Figure 5.4 and Table 5.8.

Table 5.8

LA-Descriptions and Descriptive Statistics for Scales Based on Items from Student Questionnaire

	N	Min	Max	Mean	SD	Description
eAssess	9	2.27	4.00	3.20	0.50	Perceived efficacy of computer use for the examination. Potential range between 1 and 4
eAssessP	9	3.00	3.73	3.22	0.24	Perceived efficacy of computer use for the portfolio. Potential range between 1 and 4.
Apply	8	2.00	3.00	2.58	0.35	Application of computer to various uses. Potential range between 1 and 3.
Attitude	9	2.20	3.00	2.70	0.24	Attitude towards using computers. Potential range between 1 and 3.
Confidence	8	2.50	3.00	2.77	0.18	Confidence in using computers. Potential range between 1 and 3.
Skills	8	3.00	4.00	3.63	0.38	Self-assessment of ICT skills. Potential range between 1 and 4.
SCUse	7	48	240	114	69	Estimate of time in minutes per day spent using computers at school .

An explanation of the scales was set out in Table 3.4 in Chapter Three. Results indicated that the examination and portfolio were both perceived to be appropriate instruments of assessment and relatively easy to complete. Means for *eAssess* and *eAssessP* were approximately 1.5 and 2.5 standard deviations above the mid-points respectively. *Attitude* and *confidence* scales were even more positively skewed with means approaching two and four standard deviations above the midpoints. Students had a high, self-assessed level of ICT *skills* across a range of applications (mean = 3.6, midpoint=2.5, SD=0.38). On average these students indicated using ICT for a little over 1 hour per day at school.

Student Forum

Four students (3 male 1 female) agreed to stay behind immediately after the examination to be part of the student panel. The discussion went for 4m 2s.

What did you think of the task(s) you were asked to do?

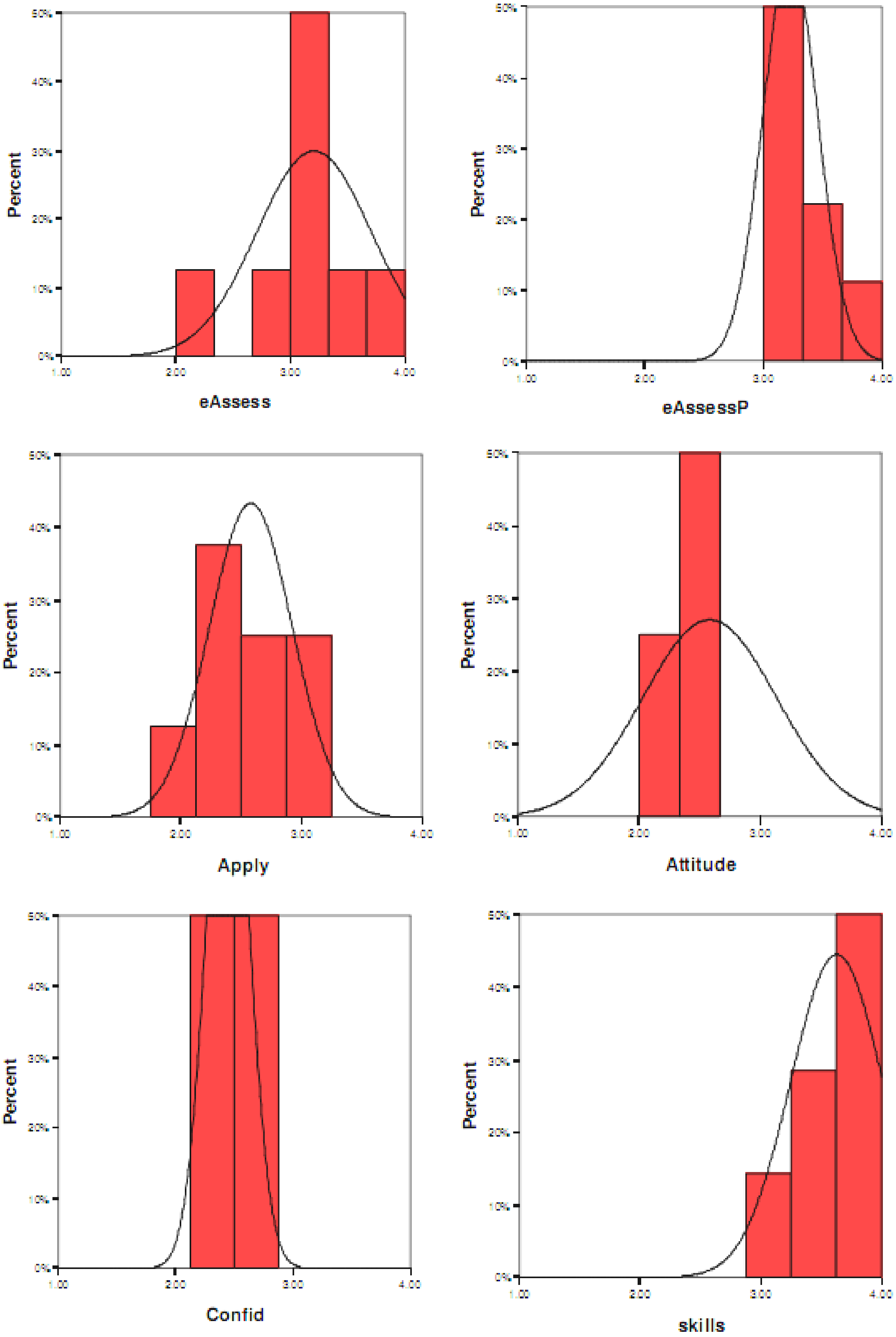


Figure 5.4 LA graphs for distribution of scores for scales on the student questionnaire.

“I thought they were straightforward” “The graph one was…pretty stupid. Any idiot can click graph on excel and copy and paste out of a text file” [laughter] “If you’re getting marked on that compared to say designing something...It’s not really…not unless you had different scaling for marks” “you could be creative...left it open...do whatever you want...manipulate images”

Were you able to do your best quality of work?
Students agreed they could do their best quality of work and that there was enough time.

“My handwriting is terrible and your hands cramp up on you over three hours…having that I could get a lot more done in the time”.

Did the computers help? “Definitely”

How much different was this to how it used to be done?
“This tops it easy…tops it so easy” “This is better because it shows your practical skills not just the planning stage” Again increased productivity of keyboarding was referred to.

What, if anything, would you like changed in future?
“The evaluation at the end [reflective questions]…it just seemed like you asked the same questions four times’ “I got four words out of a thesaurus and copied and pasted those in three or four times”.

Were there any technical problems with doing the activities?
“The microphone…at the beginning” “Maybe if you had USB headsets instead of the traditional red/green and yeah USB drivers just sort it out”.

In summary, students were very positive about the examination considering it to be a fair and comprehensive method of assessment of AIT. There were some criticisms of the tasks, the graphing with *MS Excel* was perceived as trivial and the reflective questions were also widely condemned as repetitive, though this is evidence that they were in fact misunderstood. Students suggested that more value be placed on creative tasks and that the assessment structure be more flexible.

__Email ‘Interview’ with Teacher__

What did you think of the assessment tasks overall?

> The AIT Portfolio project was very similar to the task that I intended to deliver as the production /performance task to the students. I was able to accommodate the research project quite easily into the framework of my task. The only difference was in the Evaluation section of the project. I needed to follow the framework of the sample exam provided by the Curriculum Council, so the Evaluation students submitted to me differed to that submitted to ECU. These students were preparing for an external exam in 2009 so I tried to emulate the conditions they would experience in the external exam within their tasks where possible. The Exam was well received by students. They had no problems accepting the situation of doing an exam that would not be credited as part of their AIT Unit 2B grade because it could be completed on the computer, even though it meant sitting 2 x 3hr exams and coming in at the end of the examination period. I was really delighted with their positive response to this request.

What did you think of the structure of the activities?

> The timing fitted in exceptionally well with the Course Outline and students were able to follow the instructions. The only problem I noted after the conclusion of the task was that students had been asked to submit the two artefacts and write ups and several students did not comply with this request. This may have been because they were not being formally assessed in the school-based assessment/grade. The only question I posed was in relation to the artefacts. My query was “One student wants to use a video he produced for a task that is 156MB. He is going to try to reduce the file size over the weekend using Video RA which converts to a MPEG4 or a H264 (?) file. Are you able to run the movie on Windows Media Player Classic or an iPod?” I received a positive response promptly in reply to this query.

What were the students’ reactions to the activities?

“The students all appeared to enjoy the challenges of the research project. The tasks, sub-tasks and exam were well received”.

What do you think of its potential?

> The potential of using computer hardware, software, and prepared files reflects a student's knowledge base and also their skills base in a tangible form. If we are asking our students to complete the majority of their assessments using these tools throughout the year, then surely we should in the final exam.

In summary, the teacher was positive about the intention, nature and structure of the tasks (particularly the examination) commenting on the similarity of these to her own assessments. The teacher perceived both portfolio and examination to be valid and authentic and was satisfied that these allowed students adequate opportunity to demonstrate their ability. The teacher was very much in favour of external assessment of students' practical skills. The single technical issue highlighted was resolved.

Results of Marking

The two external assessors marked five pieces of work for each student comprising the design document for the product, the product itself, two further digital artefacts, a theory section and a practical section of a three-hour examination. The marks for these were totalled and the results are shown in Table 5.9.

Table 5.9

Results for Case LA from Marking Portfolio and Exam (N=10)

St ID	Assessor marking (Total %)			Assessors		Tch (%)	Rank	
	Ass1	Ass2	Ave	Pfolio (70)	Exam (30)	Sem	Ass Ave*	Tch Sem
la101	49	34	41.5	30.0	11.5	65	45.0	7.5
la102	46	62	54.0	31.0	23.0	77	22.5	1.5
la103	17	19	18.0	12.0	6.0	66	94.0	6.0
la104	52	60	56.0	39.5	16.5	77	20.5	1.5
la105	41	55	48.0	34.5	13.5	63	30.0	9.0
la106	29	42	35.5	35.5	a	75	60.0	4.0
la107	29	46	37.5	22.5	15.0	62	55.5	10.0
la108	39	23	31.0	16.5	14.5	76	70.5	3.0
la109	34	47	40.5	22.0	18.5	65	47.5	7.5
la110	36	51	43.5	26.5	17.0	71	39.5	5.0
Mean	37.2	43.9	40.6	27.0	15.1	69.7	48.5	
SD	10.1	13.9	10.6	8.8	4.7	5.8	21.5	

*Ranking of external assessors is for all 115 students involved in the study.

a. Student did not sit exam for disciplinary reasons.

For this case no examination marks were supplied but a semester mark was. The *rank-assessor average* is the position out of the whole sample of 115 students spread over seven schools. The *rank-teacher's semester (Tch Sem)* is the rank only within the class. The mean of the average of the external assessors' marks (40.6%) was a little higher than the mean for all students involved in the study (36.9% with SD of 17.2%). There were no students in this class ranked in the top 10% of the 115 students (*Rank -Ass Ave*). The mean ranking of the external assessors was 48.5, below the mid-point of 58.

A correlation analysis was undertaken on the rankings and on the scores generated. The results are shown in Table 5.10 and Table 5.11.

Table 5.10

LA Correlation Coefficients from the Marking of Student Work (N=10)

	Assessor 1	Assessor 2	Average	Teacher Semester %
Assessor 1	1.00	0.55	0.84**	0.36
Assessor 2		1.00	0.92**	0.19
Average			1.00	0.29
Teacher Semester%				1.00

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

Table 5.11

LA Correlation Coefficients from the Ranking of Student Work (N=10)

Rank of	Assess1	Assess2	Average	Teacher Semester %
Assessor 1	1.00	0.51	0.82**	0.37
Assessor 2		1.00	0.91**	0.17
Average			1.00	0.26
Teacher Semester %				1.00

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

Correlations between the two external markers were moderate and not significant being 0.55 for the marking and 0.51 for the ranking of students. This could be explained by the small sample size of 10 participants. With no separate teacher marks for the examination, correlation between the external markers and the teacher's semester mark is even weaker at 0.37. Without investigating the content assessed it is difficult to explain why there should be such limited agreement over the ability of the same students

Conclusions

Conclusions pertaining to this case study are discussed separately for the portfolio and the examination.

Portfolio

The students, by their survey responses, and the teacher, in conversation with the researcher, indicated that they were familiar with this form of assessment, with understanding, skills and knowledge being demonstrated through performance on activities that had an obvious connection to the real world. The portfolio tasks mirrored typical class activity and were observed to be easily manageable. Collecting the student portfolios was somewhat problematic with many submissions being incomplete in whole or in part. This meant that high scores were rare. There was only limited agreement between scores awarded by the two external assessors for the portfolio and it was suggested by Assessor 1 that inability to display content when marking remotely may have been the cause. These anomalies were obvious when both sets of marks were compared and could have been rectified by remarking. The structure of the assessment tasks was favoured by both students (survey) and teacher (interview) and allowed students to fully demonstrate their ability.

Examination

With the exception of the audio recording of the students' reflection, this computer-supported production examination was implemented with no significant technical difficulties with all students able to complete the requirements in the time permitted. The students responded well to the style of examination and appeared to enjoy the practical component. Many students didn't seem to appreciate the difference between a logo and a banner or poster. Numerical data was handled well with one student suggesting that graphing of data using a chart wizard was too trivial at this level. Interestingly another student failed to produce any graphs or charts. Although agreeing that the theory section was made easier by being done on a computer, many students found the reflective questions difficult to understand and repetitive in nature. As one student said, "I got four words out of a thesaurus and copied and pasted those in three or four times." Many appeared to misunderstand the subtleties of the questions and to which part of the technology process they referred as evidenced by the many answers which were off topic. Capture of the digital content of the examination presented no difficulties with file types.

Case Study MA: Private School

The MA case study involved one teacher and a class of 12 students all of whom consented to be involved in the study. The class was conducted in the computer laboratory pictured in Figure 5.5.

Figure 5.5 MA computer laboratory

Implementation, Technologies and Issues Arising

The context for the class was Business Information Technology. The teacher implemented the tasks exactly as set out in the project documentation (see Appendix F), using the design brief for an e-commerce website as the portfolio product. The portfolio and the examination formed a part of the semester mark awarded. The researcher either met or communicated using phone and email with the teacher before the students became involved. This was to discuss the process with the teacher and to test the technologies, in particular the use of computers in the school's laboratory with sound recording and a USB flash drive. The students' computers were all less than three years old and well equipped with office and multimedia software. Observation and comment by students suggested that multi-tasking, particularly with large files, often strained system resources.

Data Collected

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 assessment tasks.

Observations of the Classes

The class was visited on four occasions: to observe students completing the portfolio tasks, to observe and invigilate the examination, to administer the student questionnaire and to conduct interviews with a small group of students about the examination.

Visit 1: Product Development (02/09/08, 11:30-12:20)

Ten students (eight boys and two girls) were observed working for about one hour on portfolio component 1 the *Miss Shoppe* e-commerce website. Two students were away at Structured Workplace Learning. They had started the previous week on the *Investigation* and *Design* sections of the project. All but one were observed to be working on the investigation or design, completing reports, concept maps, storyboards, graphic layouts and looking at examples of websites (e.g. *RipCurl* and *Billabong*). Some were working on paper, others in *MS Word*, *Adobe Fireworks* or *Dreamweaver*. One boy had begun production of graphics in *Adobe Fireworks* and another claimed to have done the *Design* at home, on paper, and was now going back to do the *Investigation* (he said he would modify his design if he found anything useful, otherwise he wanted to focus on his own ideas). Another boy indicated that he didn't like using paper. The teacher had specified that website production should be in *Dreamweaver* with graphics done with *Adobe Fireworks* but left other choices up to the students themselves.

The Teacher instructed the class for a few minutes using the whiteboard, focussing on the important features of a storyboard. The teacher gave verbal feedback to some students on their designs. Two girls were doing graphic designs of their main interface using *Adobe Fireworks*. About half the boys appeared to make very limited progress during the session.

Visit 2: Development of Design Process Document (21/10/08, 11:30-12:20)

Eleven students (nine boys and two girls) were observed working for about one hour on portfolio component 2, the *Design Process Document*. Two students were away at Structured Workplace Learning. A new student had recently arrived from UK but would not sit the exam. The researcher was introduced by the teacher and thanked the students for their participation. The purpose and importance of the project was again briefly explained. The exam structure was also described and the importance of the questionnaire and student forum was also highlighted. The class was working on the *Design Process Document* using the template supplied. The completed work was due in the next day (Wednesday) for most with an extension for some to Friday. Students were observed to have digitised their design sketches and added these to the template filling out the prompts in the *investigate* /*produce* / *evaluate* sections. The teacher was asked by the students about peer evaluation and responded that there was no time for this now. He added that, since time was short, a self-evaluation of just the strengths of the product and its perceived weaknesses with perhaps some suggestions for improvement would suffice. Students were happy to show and tell their websites. These were made from scratch in *Adobe Dreamweaver* (i.e. not developed from templates) with *Adobe Fireworks* being used for banners, graphics, navigation buttons etc. To the researcher, some looked quite good but most were well short of publishable quality. Students obviously knew their way around the applications at their disposal.

USB memory and headset/microphones were checked on this visit. All were working with audio recording possible at very low volumes with good quality and this would help manageability of the exam. Students used *Windows Sound Recorder* (XP version) though appeared unaware that this was limited to a minute capacity. *Audacity* was also available and this would be suitable for extended recording.

The exam was scheduled for 3 hours continuously in week 6 of term (17th-21st November) and a morning time slot had been requested. The questionnaires and student forum would have to be done on that day as these students would be on study leave and then away until the following year as soon as exams finished. The teacher was well organised having already collated the portfolio files according to student codes and these could conveniently be picked up on the day of the exam. The researcher explained about the order of sections and the problem this had created at RA and it was agreed that doing Part B (the practical component) first was preferable. The idea of a sound test before the exam started was also discussed. The idea of slipping in the questionnaire to those who finished the written reflection early was also broached. There would still potentially be invigilating problems with timing and restrictions over sections but this was a small group and would be well spread out in room with spare computer capacity. The exam paper was to be re-written in two separate parts and these would be passed on to the teacher for printing. The teacher was reminded of the practice exam which had been developed and this would be run through over the following couple of weeks. The teacher said he needed to cover some of the *Excel* and business context material from the 2B course but had every confidence this would go well.

Visit 3: Examination (20/11/08, 7:00-12:00)

Thirteen students were present for the examination in the same computer lab as previous visits with one student not sitting the full examination leaving after 2 hours. The exam set up took about 20 minutes after which students entered and were given a short briefing explaining the structure of the paper and the sections timings. Reading time began after an audio test. No issues were evident. Students were able to browse the photographs and data during this time. Three students' reported that files on the USB flash memory were corrupted and this problem was quickly remedied by reimaging the device from a spare. This was quick to fix and involved minimal disturbance.

The exam started at 8:20. All students were observed to plan on paper and this made invigilation simpler as it allowed enforcement of the 15 minutes minimum planning time during which work on the computer was not allowed. Students, quite sensibly, appeared to be using the photographs to stimulate their design ideas. It was noticed that the marks for each section of the exam didn't reflect the time allocated to each but it was too late to change anything. At 8:35, planning time ended but most students kept on planning; this was sensible really in terms of mark allocation. The students appeared to know how to design and annotate using storyboards. For logo creation students went mainly with *Adobe Fireworks*. Of the logos observed many incorporated photographs or elements and ideas from photographs and were not really logos. After 30 minutes, student MA108 was still designing on paper. Student MA111 was cropping out a turtle for the logo. Logos were observed to evolve from the supplied photographs with writing added, suggesting that students did not fully understand what a logo was. The data text file for student MA110 was missing but this was quickly fixed. A few students began graphing with *Excel*. Student MA108 made a logo from scratch using *Adobe Fireworks* but it was yet another turtle. Student MA112 spent a long time on the logo. She sampled parts of the photos to get her colour scheme and it would be interesting to hear her explanation of this method. After about 60 minutes: Student MA102 was observed to be using *MS Publisher* for the brochure. Student MA106 had a few problems with the chart wizard in *MS Excel*.

The audio recording needed more time; 5 minutes was too short. It was difficult to strictly enforce the timing and this section overran a little before change over to the reflective questions. These were on a separate paper. A few students asked questions on wording and it was surprising that students felt that this was permissible in an exam. Students' keyboarding was a mixture of *touch* typing and *hunt and peck*. The reflective questions were again completed within the one hour time allocated allowing students to begin the questionnaire.

Visit 4: Survey and Forum (20/11/08, 11:20-12:00)

On completion of the examination students were presented with a questionnaire. A group of four students agreed to take part in a student forum. They were presented with the same set of questions as other cases but follow up questions differed depending on responses.

Survey of Students

Ten of the twelve students completed the questionnaire immediately on completion of the performance examination. The minimum, maximum, mean and population mean (mean for all 115 students) were calculated for each closed response item using SPSS (refer to Appendix K).

Items Concerning the Portfolio

Students indicated that they had some previous experience of completing a portfolio on computer. Responses to the item *how often have you done a portfolio on computer before?* produced a mean of 2.7 on a scale of 1 (Lots of experience) to 4 (No experience). They claimed that it was *quick* and *easy* to develop their *ideas* and complete the portfolio *product* (items P2(a)-P2(c) and P2g)) with means for these survey items ranging from 1.5 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree). There was also a strong positive response to the survey items P2(f) t*he computer was good for showing my skills in the portfolio* and item p2j *overall I was able to show what I can do* with means of 1.5 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students' preference for computer-based portfolios was overwhelming with responses to items such as i*t was better doing the portfolio on a computer than on paper* and *overall the computer is good for portfolios* yielding means of 1.4 and 1.3 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to population means of 1.6 for the same items.

Items Concerning the Examination

Students indicated that they had little (6 participants) or no experience (4 participants) of completing an examination on computer. The mean of responses to the item *how much more time would you need to get used to it?* was 2.7 on a scale1 of 1 (Lots of time) to 4 (No time) slightly below the population mean of 2.8. Students were very positive about the examination, with responses to questionnaire statements containing *quick*, *easy* and *good* producing means ranging from 1.4 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree). The item *overall I was able to show what I could do in the exam* produced a mean of 1.6 compared to the population mean of 1.9. Students' preference for a computer-based examination was strongly supported with the statement *it was better doing the exam on computer than on paper* producing a mean of 1.7 on a scale of 1 (strongly agree) to 4 (strongly disagree).

Items Concerning Computer Use

These students were regular and extensive users of new technologies with all using a computer with broadband at home on a daily basis for communication or educational purposes. A full range of home entertainment systems was indicated and mobile 'phone ownership was almost total. *Attitude* to computers was definitely positive, for example all students agreed with the statement *computers are good for the world*, and there was widespread support for statements *such as I enjoy using computers at school* and *I feel confident working with computers* with responses ranging from 1.0 to 1.2 on a scale of 1 (Yes) to 3 (No). Students were unsure of whether or not they could learn to program a computer (mean 2.0). Student's self-assessment of their proficiency with applications was also high with the exception of databases and spreadsheets. Means for these items were 2.3 and 2.7 compared to 3.4 for word processing, 3.6 for image editing and 3.9 for internet research on a scale of 1 (Low ability) to 4 (High ability). These results were largely in line with population means.

Open-ended Items

Responses to the open-response items were tabulated to assist in drawing out themes. Comments were varied and a sample of these is shown in Table 5.12.

Table 5.12

MA Summary of Students' Responses to Four Open Ended Questions on Survey Questionnaire

The two best things …	The two worst things …
Portfolio 6 students made reference to innovation e.g. "learn how to produce new things" and "gave me new skills" and "taught me how to create a website". Other responses alluded to general *ease* e.g. "easy to make changes" and "don't have lots of paper to worry about" and "don't have to write it up".	Portfolio 7 responses mentioned time as a problem for example "Not enough time" and "the rate at which we were expected to work" and "took a lot of time". There were other themes. Some other responses were "you might not save work (lose all work" and "the theory side" and "couldn't pick up what I have to create".
Examination Easy (9 students referred to the superiority, ease, speed and accuracy of keyboarding over writing e.g. "It was a lot easier to use the computer than a pen" and "Hands don't get sore". Creative (1 student mentioned the creative advantage of the computer, "More creative than just writing"	Examination Multiple Designs (3 students mentioned their dislike of having to do two designs e.g. "Doing two designs" System failure (3 students alluded to the ever present 'danger' of something catastrophic going wrong e.g." Computers can always have a problem "and "concern that hardware may fail". Noise of others typing was also cited by 2 students.

There were four open-ended questions on the student questionnaire. Students were asked to list the two best things about doing the portfolio and examination by computer and to list the two worst things about doing the portfolio and examination by computer. Generally students considered that using computers made it easier, particularly by removing handwriting and also provided a better environment within which they could use their skills and demonstrate their ideas. The main worst things were having to do multiple designs and general worries about system failure, processing delays caused by multitasking applications and a prevailing fear that work might be lost.

Questionnaire Scales

Some of the results of an analysis of the seven scales derived from combining selected items from the questionnaire are shown in Figure 5.6 and Table 5.13. An explanation of the scales is set out in Table 3.4 in Chapter Three.

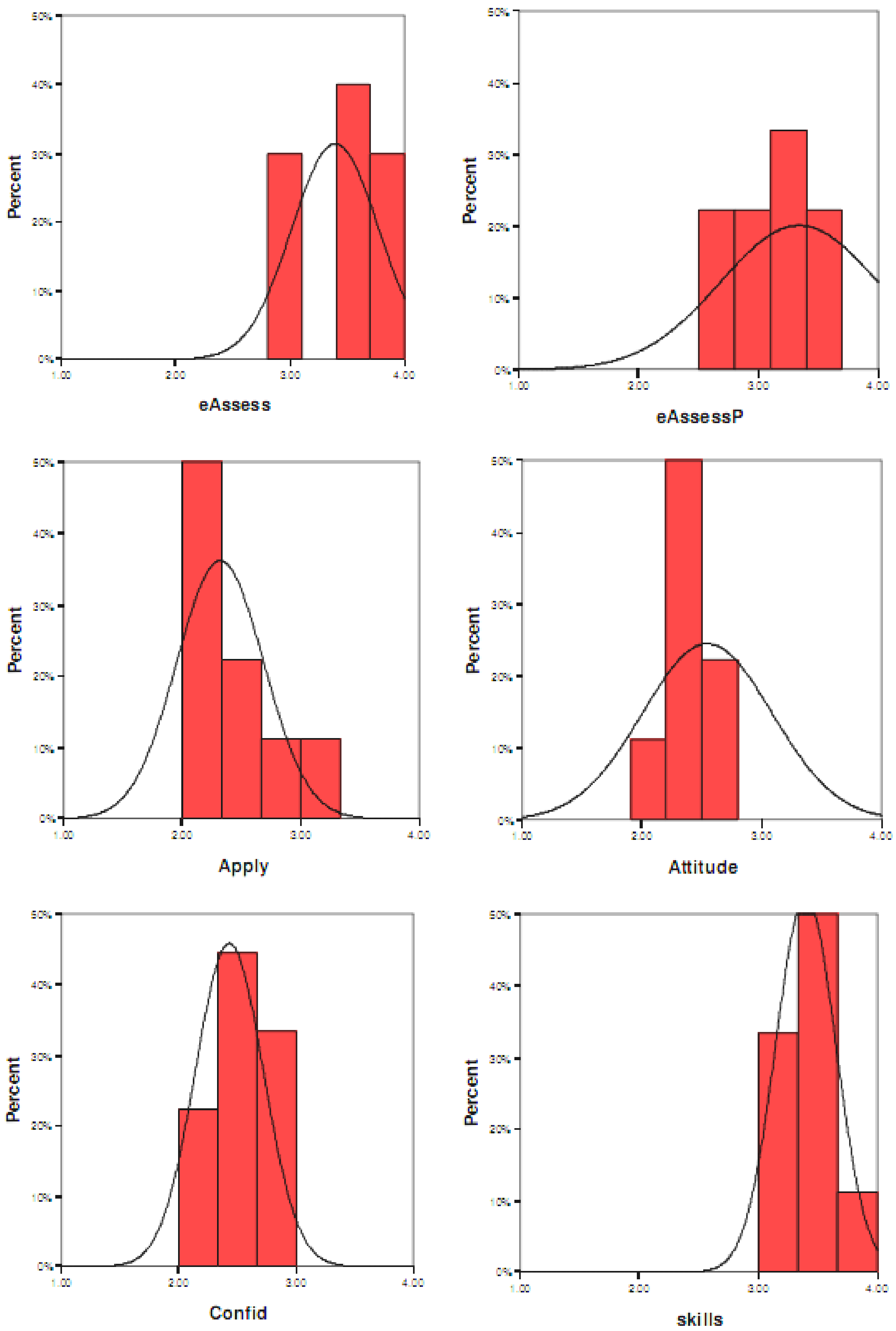


Figure 5.6 MA-graphs for the distribution of scores for scales on the student questionnaire

Students indicated that the examination and portfolio were both appropriate assessment instruments which were relatively easy to complete with means for the scales *eAssess* and *eAssessP* were approximately 1.5 and 3.5 standard deviations above the mid-points. Students' *attitude* and *confidence* were also highly positive with means on these scales being between two and three standard deviations above the mid-points

Table 5.13

MA-Descriptions and Descriptive Statistics for Scales Based on Items from Student Questionnaire

	N	Min	Max	Mean	SD	Description
eAssess	10	2.82	3.82	3.69	0.78	Perceived efficacy of computer use for the examination. Potential range between 1 & 4
eAssessP	9	2.73	3.64	3.64	0.32	Perceived efficacy of computer use for the portfolio. Potential range between 1 & 4.
Apply	9	2.00	3.00	2.31	0.37	Application of computer to various. Potential range between 1 and 3.
Attitude	9	2.40	2.80	2.64	0.17	Attitude towards using. Computers. Potential range between 1 and 3.
Confidence	9	2.00	3.00	2.69	0.32	Confidence in using computers. Potential range between 1 and 3.
Skills	9	3.00	3.73	3.39	0.25	Self-assessment of ICT skills. Potential range between 1 and 4.
SCUse	9	36.00	333	99	93	Estimate of time in minutes per day spent using computers at school.

A high level of self-assessed ICT skills across a range of applications was indicated. On average students reported using ICT (*SCUse* in Table 5.13) for a little over 1½ hours per day at school.

Student Forum

Four students remained behind to form a focus group. The discussion ran for 5m 12s.
What did you think of the task(s) you were asked to do?
It was generally agreed that the tasks were "fair" and "Quite easy really" though the time allowed was problematic: "a bit rushed".

Were you able to do your best quality of work?
Students said they were able to do work of good quality given the time.

Did the computers help?

“Yes. definitely easier”. “A lot quicker to type than to write”. “Mistakes are easy to fix.”

How much different was this to how it used to be done?
“This is a much better” “A lot simpler” “because you can type - you don’t have to worry about mistakes.” “Practical lets you show what you can do…I’m not great at theory.”
All agreed they were able to produce a lot more in the given time.
What, if anything, would you like changed in future?
“Allow more time”. All concurred with this sentiment. “Or just make it less to do”.

Were there any technical problems with doing the activities?
There were no major issues here with sound recording. Some of the files on the memory devices were corrupted but this was simple and quick to put right.
Were there any other problems with the activities?
The wording in the reflection questions appeared to be confusing. “It took me a while to work out what was meant by it.” Students would like simpler, clearer instructions and felt the language used was a little too advanced for them.

Any other thoughts or suggestions for developing the use of digital forms of assessment?
“Having faster computers …it took a while to load some programs” and this was a constant worry for students.

In summary, students felt the computer-based examination was a fair form of assessment allowing them to show what they could do and with compelling advantages over paper based assessment. Students also indicated that there was perhaps a little too much to do in the time allowed for the practical section and were fazed by the language used in the theory section. On balance students perceived the examination highly positively.

Email ‘Interview’ with Teacher

What did you think of the assessment tasks overall?

> The tasks were good but I thought my students may have performed a little better with the additional pointers and scaffolding but in reality, they still performed at

about the same level. Could not see much evidence of performance differing between the tasks completed at the start of the year and the ECU Tasks. The practical component of the exam was completed well but the theory section was not rigorous enough. The marking guides were quite limiting and did not cover enough of the task. They needed to cover more of the process than they did.

What did you think of the structure of the activities?

The structure of the activities was good. The timing and structure was fine with the portfolio and instructions were quite clear. There were some timing issues with the exam. Students seemed rushed. The response section of the practical exam, students wrote a response which was marked and then recorded the written response which was marked separately. Not sure if that was the intention but the marking ended up being a lot for that small part of the work.

What were the students' reactions to the activities?

The students enjoyed the portfolio task and had very little issues or problems. Students liked the idea of the practical exam and I expected better performance then was actually achieved. Not sure if that is my expectations or the marking scheme or something else?

What do you think of its potential?

The potential of a practical exam is good however it would need to be a little more flexible such as a website/pages, logo, brochure etc. as the range of practical assessments in the course varies a lot. The portfolio I can see will be very time consuming for students to create however it would be a better way of assessing than just a paper exam.

What did you think of the quality of work produced by your students for these tasks?

The quality of some of the portfolios were good and in some cases better than expected. The quality of the exam I thought was quite poor given that it was practical and not paper-based.

Were you surprised by the performance/attitude of any students?

“As I have already stated I was disappointed with the students’ exam. I was pleased with the quality of most students portfolio work”.

What was the general feedback from students?
“Students liked the idea of the practical exam but performance was still similar to other tasks they had completed in the course. The portfolio reflection questions were not received well at all particularly by male students”.

Were there any technical problems with implementing the activities?
“Very little technical problems”.

Were there any other problems with implementing the activities?
“Some items were repeated a few times and this was not received well”.

Any other thoughts or suggestions for developing the use of digital forms of assessment?

> A better, more comprehensive marking scheme that reflects all work that students complete e.g., marks out of 3 does not give enough scope for assessing some sections of the practical work or exam. A better fit with the entire syllabus as there was too much work to cover the syllabus in its entirety and complete the portfolio and other tasks as well. Assessments could actually reflect some of the theoretical sections of the course e.g. students completed their portfolio as a website covering a series of social implications and trends section of the course.

In summary, although positive about the concept of portfolio and computer-based assessment, the teacher made several constructive comments about the structure of tasks themselves. The audio recording section was seen as a duplication of the written reflection and the time allocated to the practical section was perceived as too short. The teacher commented adversely on the mark allocations in the examination and suggested having students present some of their responses to the theoretical sections of the course in the form of practical activities, for example as websites.

Results of Marking

The two external assessors marked five pieces of work for each student comprising the design document for the product, the product itself, two further digital artefacts, a theory section and a practical section of a three-hour examination. The marks for these were totalled and the results for each student are shown in the Table 5.14.

Table 5.14

Results for Case MA from Marking Portfolio and Exam (N=12)

St ID	Assessor marking (Total %)			Assessors		Teacher marking (%)		Rank of		
	Ass1	Ass2	Ave	Pfolio (70)	Exam (30)	Exam	Sem	Ass Ave*	Tch Exam	Tch Sem
ma101	53	68	60.5	43.5	17.0	70	69	13.0	2.0	2.0
ma102	41	36	38.5	25.5	13.0	37	57	53.0	6.0	6.0
ma103	32	19	25.5	13.0	12.5	28	35	83.5	12.0	12.0
ma104	65	43	54.0	37.0	17.0	53	55	22.5	7.0	7.0
ma105	42	43	42.5	32.0	10.5	46	46	44.0	9.0	9.0
ma106	32	34	33.0	23.0	10.0	37	40	64.5	10.0	10.0
ma107	69	62	65.5	45.0	20.5	74	73	8.0	1.0	1.0
ma108	62	76	69.0	46.5	22.5	64	61	4.0	4.0	4.0
ma109	59	53	56.0	41.5	14.5	70	69	20.5	3.0	3.0
ma110	39	41	40.0	29.0	11.0	38	47	49.0	8.0	8.0
ma111	41	34	37.5	22.5	15.0	34	38	55.5	11.0	11.0
ma112	53	40	46.5	29.0	17.5	57	59	34.5	5.0	5.0
Mean	49.0	45.8	47.4	32.3	15.1	50.7	54.1	37.7		
SD	12.8	16.2	13.6	10.5	4.0	16.1	12.9	24.7		

*Ranking of external assessors is for all 115 students involved in the study.

The teacher provided a separate score for the examination and one for the semester which included the portfolio and the examination. The rank *Ass Ave* is the position out of the whole sample of 115 students spread over seven schools. The *rank teacher's semester (Tch Sem)* is the position of the mark awarded by the teacher compared to the marks awarded by all other teachers in the sample. The mean of the average of the external assessors' marks (47.4%) was statistically significantly higher than the mean for all students involved in the study (36.9% with SD of 17.2%). There were two students in this class ranked in the top 10% of all students *(Rank Assess Ave).* The mean ranking of the external assessors was 37.7, which was below the mid-point of 58. Overall this class performed well compared with the other classes.

A correlation analysis was done on the rankings and on the scores generated. The results are shown in the Table 5.15 and 5.16.

Table 5.15

MA Correlation Coefficients from the Marking of Student Work (N=12)

	Assessor 1	Assessor 2	Average	Teacher Exam %	Teacher Sem %
Assessor 1	1.00	0.75**	0.92**	0.87**	0.83**
Assessor 2		1.00	0.95**	0.86**	0.79**
Average			1.00	0.93**	0.87**
Teacher Exam %				1.00	0.93**
Teacher Semester %					1.00

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

Table 5.16

MA Correlation Coefficients from the Ranking of Student Work (N=12)

Rank of	Assessor 1	Assessor 2	Average	Teacher Exam %	Teacher Sem %
Assessor 1	1.00	0.80**	0.94**	0.89**	0.83**
Assessor 2		1.00	0.95**	0.86**	0.81**
Average			1.00	0.92**	0.86**
Teacher Exam %				1.00	0.94**
Teacher Semester %					1.00

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

Correlation between the external markers was strong and significant for scores awarded ($r=0.75$, $p<0.01$) and for the ranking of students ($r=0.80$, $p<0.01$). When compared with the average of the two assessors, the correlations on the teacher's scores for the examination and semester were very strong. ($r=0.93$, $p<0.01$ and $r=0.87$, $p<0.01$ respectively).The rank of teacher marks for the examination and semester also correlates strongly with the rank of the average of the two assessors. ($r=0.92$, $p<0.01$ and r=0 .86, $p<0.01$ respectively).

Conclusions

Conclusions pertaining to this case study are discussed separately for the portfolio and the examination.

Portfolio

Submissions of student work were mostly complete allowing some high marks to be awarded. Two of the 12 students were ranked in the top 8 of all 115 students in the population by the external assessors. It was notable that in this case the portfolio task was observed to be delivered by the teacher exactly as intended and although no separate scores were provided, the portfolio was included in a full set of teacher marks for the semester. Of the seven cases studied, this was the closest to full compliance by the teacher. The teacher and students expressed positive sentiments towards the nature of the tasks and some imaginative and original work was produced. The students, by their survey responses, and the teacher, in conversation with the researcher, indicated that they were quite familiar with this form of assessment, with understanding, skills and knowledge being demonstrated through performance on real world activities. The portfolio tasks mirrored typical class activity and were observed to be easily manageable.

Examination

The computer-supported production examination was implemented with no significant technical difficulties with all students able to complete the requirements of the exam in the time permitted. Students completed the reflective questions well inside the time allowed though analysis of the responses showed a widespread misunderstanding of the intent and depth required. The students responded well to the style of examination and appeared to enjoy the practical component. Many students didn’t seem to appreciate the difference between a logo and a banner or poster. Numerical data was handled poorly by a couple of students. Although agreeing that the theory section was made easier by being done on a computer, many students found the reflective questions difficult to understand and repetitive in nature. Many appeared to misunderstand the subtleties of the questions and to which part of the technology process they referred. “It took me a while to work out what was meant by it.” Students would like simpler, clearer instructions and felt the language used was a little too advanced for them. Many answers were off topic.

Case Study RA: Public School

The RA case study involved one teacher and a class of 14 year 12 students studying AIT at *Stage 2* in the context of Business Information Technology. The teacher implemented the tasks as supplied using the design brief for the e-commerce website as the product for the portfolio (see Appendix F). The portfolio and the examination formed the basis of the semester mark awarded but the exact make-up of the mark was unknown.

Implementation, Technologies and Issues Arising

The Researcher either met or communicated using phone and email with the teacher before the students became involved. This was to discuss the process with the teacher and to test the technologies, in particular the use of computers in the school's laboratory with sound recording and a USB flash drive. The class was conducted in the computer laboratory pictured in Figure 5.7.

Figure 5.7 RA computer laboratory

The students' computers were networked to a central server which delivered the application software. There was an abundance of software available and students often had a choice of application with which to work. The network, in combination with the desktop hardware made for frequent delays in processing larger files. Sometimes, computers would stop responding and have to be rebooted. Observation and comment by students suggested that multi-tasking, particularly with large files, often strained system resources. Comments from the network manager suggested that the delivery system was under-resourced and an upgrade was planned for later in the year.

Data Collected

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 assessment tasks.

Observations of the Classes

The class was visited on four occasions: to observe students completing the portfolio tasks, to prepare for the examination, to observe and invigilate the examination, to administer the student questionnaire and to conduct interviews with a small group of students about the examination.

Visit 1: Product Development (30/07/08, 11:30-12:20)

Fourteen students (four girls) were present in the computer laboratory. The room contained 24 computers, in two columns of facing seats and one column wall facing, and were well spaced with system units located on the bench top in between making it difficult for students to observe the work of others. There was sufficient writing space. The class was addressed to explain the purpose and value of the project, the requirement of consent forms and the fact that the work was part of the school based assessment schedule. The class had already spent a week on the task using a four-page hand-out taken from the document *AIT Project for Portfolio Assessment* and were following the documentation supplied with the project without modification.

All students were observed to be keeping a journal of each session and had done so for the year. The teacher described the wide variety of abilities in the class stating that about five students were of capable of tertiary education entrance. Some network issues were immediately evident; for example slow logging in. All students started work quite quickly, mainly using *MS Word*. Some worked with *Paint*, *Internet Explorer* and *FrontPage*. A few had started or were starting design using paper based brainstorms, others were developing storyboards containing mock-ups of webpage designs. One girl was creating a background graphic and three others were creating logos and banners for their website. One boy was clearly off-task. About half appeared to have done most of the research and some were taking creative directions (e.g. *Shoddy Shoppers Store*). All seemed to know what to do and stated when questioned that this was a familiar type of task and process. Two students created multiple page designs in *Paint*. Each student worked independently but bounced ideas off each other and the teacher, who asked some students to report on what they had done so far saying that was practice for the oral part of examination. It would have been ideal to have stopped the class 15 minutes from the end so that students could have explained what they had done and why, thus providing information on their design process and technique.

Visit 2: Development of Design Process Document (27/08/08, 11:30-12:20)

Only 12 students were present because of a field trip. The class had only just begun work on the process document and the lesson began with an explanation by the teacher and researcher of the requirements of the document. The design process template document was displayed on the data projector. The examination was also discussed and the students were made aware that they would be completing a practice examination beforehand. Students seemed clear on the requirements of the *Design Process Document* and began or continued working. Some students had the *AIT Project for Portfolio Assessment* document open and were word processing the design document. Some were still working on the portfolio product, the e-commerce website, using *MS FrontPage* or *Adobe Dreamweaver*. These students were collecting pictures from *Google* principally images of fashion items to populate their pages and *Google Maps* to show the location of their store. One boy was using non-school software running from his own USB flash drive to make his pages by typing a combination of *Java Script* and *HTML*. A couple of students hadn't started the portfolio task at all yet and were being coached by others who had. These students went straight to *Dreamweaver* without any investigation or design. The class appeared fully engaged throughout. As these were year 12 students with only four weeks left, the class teacher informed them that they would be working towards the practical exam when the portfolio was concluded.

After the lesson, several hardware tests were completed by the Network Manager in preparation for the examination. The current student image had the sound card disabled and since the image could not be edited it was agreed to try to install the sound card and drivers machine by machine. The students had *Audacity* installed or could use *Windows Sound Recorder* though this has only a 60 second recording capacity (the examination required the recording of a 30 s sound clip). A headset and microphone were left behind in order to facilitate further sound tests. A server upgrade was planned for the break and with the examination being scheduled for the break too there might be problems ahead. However the Network Manager confirmed that in this event, the examination would take priority and no major system changes would be done until its conclusion.

Visit 3 Examination (02/10/08, 8:00-12:00)

Fourteen students were present in the computer laboratory for the examination which was conducted over a continuous period of three hours with ten minutes reading time. This was the first of the seven schools to take the examination and an extra visit had been made the previous day to make sure everything was ready. The examination began with ten minutes reading time and students then commenced section A, the reflective questions. Some students finished after 20 minutes and at least two became engaged in computer games while waiting for the start of Section B after one hour. After 50m it was apparent that all students had completed Section A. Invigilation became difficult as students waited to be allowed to begin Section B. It was impossible to know if students were merely browsing data and image files or were in fact engaged on the next task. The practical section began with much relief all round.

Student RA112 used the supplied photographs (which were intended to be manipulated and used in the brochure) in her design ideas and simply added some text to make a logo. Student RA104 imported a graphic of a palm tree to use in his logo. Student RA118 used *Photoshop* to manipulate a photo for the logo. Student RA110 found that *Photoshop* would not open on her computer and had to be moved to reserve work station. Student RA104 used a *Publisher* template for his brochure. RA111 used *Paint*, with a landscape page divided into three sections, to make a tri-fold brochure. Student RA110 appeared confused between *design* for the logo and the actual logo. His design (done using the template supplied in *MS Word*) became his logo.

After 1h 50m, students RA111and RA112 indicated that they had finished. The student questionnaire was given to keep them busy. They finished this quickly and started playing computer games. After 1h 55m, student RA104 reported that he could not save his *Publisher* file as a *PDF*. This option was not available in the version of *Publisher* deployed and there was no other PDF converter installed. A General announcement had to be made to save files in their native format if they could not be converted to the form required by the exam. After 2h 10m, more students were observed to have finished and were playing games. A few whispers were audible and students had to be reminded that examination conditions were in force. After 2h 35m, student RA116 *lost the desktop* and could not open any application software. The Network Manager was called while the student was moved to a reserve workstation.

On 2h 45m, a general announcement was made that students should be preparing for their audio reflection by completing the *MS Word* template. Student RA106 requested help for a computer hang while saving a *Publisher* file. The problem was fixed, with assistance from the invigilator, by closing some of the many open files using *task manager*.

After 2h 50m a further general reminder was announced about the audio recording and on 2h 55m recording of audio reflection commenced. There was some laughter/embarrassment and it quickly became apparent that several systems were not allowing students to record. (As this was the first implementation, no sound test had been undertaken prior to the exam). The failure came as a surprise as assurances had been received the previous day that everything was in order. It was too late to attempt troubleshooting and a general announcement was made to complete this section if possible and otherwise skip it. There was a lot of disturbance and the examination ended in disarray.

This was the first time the examination had been attempted and several problems immediately became apparent. The first of these concerned the intended section split and timing, with one hour of reflective questions, *Section A*, preceding two hours of practical, *Section B*. Firstly, students were permitted to make reference to their home directory in answering the reflective questions and were able to draw upon work of a similar nature already contained there, if they chose to do so. *Section A* was completed by most students well within the 60 minutes allocated for it. Since neither section presented in any distinguishable form to the invigilator, both sections being done on the computer, there was no effective way to determine which section a student was attempting and so to prevent a student from continuing to section B before the 60 minutes was up. For this reason it may be more manageable to reverse the sections, giving out the *Section B* as a separate paper after 60 minutes had elapsed or allowing students to manage their own time with a recommendation that they spend 60 minutes on *Section A*. A similar problem presented with the planning time for *Section B*. Here the intention was that the students spend at least 15 minutes developing design ideas before proceeding with development. The restriction was again impossible to enforce effectively. The second problem centred on the audio recording. Since examination conditions prohibited talking, students could not test their audio system prior to the commencement of sound recording during the final five minutes of the examination. It would be a good idea if, prior to the commencement of the examination, students were allowed to test the capability of their audio recording system. Thirdly, it seemed advisable to have at least a couple of spare workstations available in the event of computer malfunction.

Visit 4 Survey and Forum (20/11/08, 7:00-12:00)

On completion of the exam, students completed a questionnaire, for most, having finished early, this occurred during the 3 hour examination. After a lot of persuasion, a group of four students (3 male 1 female) agreed to remain behind to take part in the discussion forum. They were presented with the standard set of questions (see Appendix I) but follow up questions differed depending on responses.

Survey of Students

Fourteen students completed the survey questionnaire consisting of 70 closed response items and four open-response items. The minimum, maximum, mean and population mean (mean for all 115 students across the seven cases) were calculated for each closed response item using SPSS (refer to Appendix K).

Items Concerning the Portfolio

Nine out of 12 students indicated that they had lots or some previous experience of completing a portfolio on computer with responses to the item *how often have you done a portfolio on computer before?* producing a mean of 2.3 on a scale of 1 (Lots of experience) to 4 (No experience). Students also agreed or strongly agreed with the statements in survey items P2(a)-P2(h) which asserted that the computer was *quick*, *easy* and *good* for developing their *ideas* and completing the portfolio *product* with means for these items being in line with the population, ranging from 1.6 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students’ preference for computer-based portfolios was strong with responses to items such as *it was better doing the portfolio on a computer than on paper* and *overall the computer is good for portfolios* yielding means of 1.6 and 1.7 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to population means of 1.6 for the same items.

Items Concerning the Examination

Eleven of the 14 students indicated that they had no previous experience of completing an examination on computer. Responses to the item *how much more time would you need to get used to it?* varied, with a mean of 2.7 on a scale of 1 (Lots of time) to 4 (No time). Students were positive about using the computer for the examination, agreeing or strongly agreeing with statements suggesting it made things *quick* and *easy* and was *good* for creating the logo, graphs and brochure; means for these items ranged from 1.4 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree). The item *overall I was able to show what I could do in the exam* produced a mean of 1.8 compared to the population mean of 1.9. There was one slight anomaly in the response to item E2(h) *the steps of the exam helped me to develop my ideas* to which six students disagreed (the mean for this item being 2.3 compared to 1.9 for the population). Overall, students' preference for a computer-based examination was strongly supported with the statement *it was better doing the exam on computer than on paper*- producing a mean of 1.6 on a scale of 1 (strongly agree) to 4 (strongly disagree).

Items Concerning Computer Use

Students indicated that they were regular and extensive users of new technologies with all using a computer with internet access (11 out of 14 had broadband) at home on a daily basis for communication or educational purposes. A full range of home entertainment systems were reported and all but one student owned a mobile 'phone. Attitudes to using computers were positive and students indicated strong self-belief, for example 11 out of 14 responding affirmatively to the statements *I feel confident working with computers* (mean 1.2) and *I am good at using computers* (mean1.2) on a scale of 1 (Yes) to 3(No). No student disagreed with the statement *computers are good for the world*, though eight students responded *sometimes* and there was support for statements such as *I like to use a computer at home to do school work* with a mean of 1.2 on a scale of 1 (Yes) to 3 (No). Students self-assessment of their proficiency with application software was also high with the exception of databases and web authoring. Means for these items were 2.7 and 2.9 compared to 3.9 for word processing, 3.4 for image editing and 3.8 for internet research on a scale of 1 (Low ability) to 4 (High ability). These results were largely in line with population means.

Open-ended Items

A summary of responses to the four open-ended questions are shown in Table 5.17.

Table 5.17

RA Summary of Students' Responses to Four Open Ended Questions on Survey Questionnaire

The two best things …	The two worst things …
Portfolio	Portfolio
Responses were varied. Five students made reference to the "easy". (E.g. "Questions were easy to follow" and "Set out which made it easier"). The other main theme was creative/ innovative mentioned by six students. (e.g. "Learning new things like how to make a webpage" and "I learned something new" and "Being creative").	Almost half the responses were blank. The only major theme was the reference made to hardware/software issues by six students. For example, "Waiting for things to load" and "Buggy computers" and "The availability of good computers" and "Not all programs available".
Examination	Examination
Practical. Six students made reference to the practical nature of the exam. (e.g. "The fact that we actually get to make our products not just plan it " and "The fact we got to do a computer course exam on a computer"). Keyboarding. Four students made reference to their preference over handwriting. (e.g. "I can type both quicker and obviously neater on a comp").	Systems. Nine students faulted the hardware, delays and risk of losing work. E.g. "The computers being slow and freezing at every chance" and "Something could go wrong with your computer causing information loss". Distraction. Three students mentioned being distracted. E.g. "The constant sound of typing" and "Easy to get distracted"

Generally students were positive about using computers which made for quicker and easier working and allowed them to create rather than just design. The main worst things were a concern that the computers were liable to hang and could crash and their work might be lost.

Questionnaire Scales

Seven scales were derived from combining selected items from the questionnaire with results shown in the Figure 5.8 and Table 5.18. An explanation of the scales is set out in Table 3.4 in Chapter Three. Survey responses indicated that students perceived both the examination and the portfolio to be appropriate assessment instruments and to be relatively easy to complete with means *eAssess* and *eAssessP* approximately 1.5 standard deviations above the mid-points. *Attitude* and *Confidence* scales were also positively skewed with means approximately one standard deviation above the midpoints. Students had a high, self-assessed level of ICT skills across a range of applications (mean=3.3, midpoint=2.5, SD=0.48). On average they indicated using ICT for a little over one hour per day at school. These results are in keeping with other groups though the means for *Apply* and *Attitude* are somewhat lower.

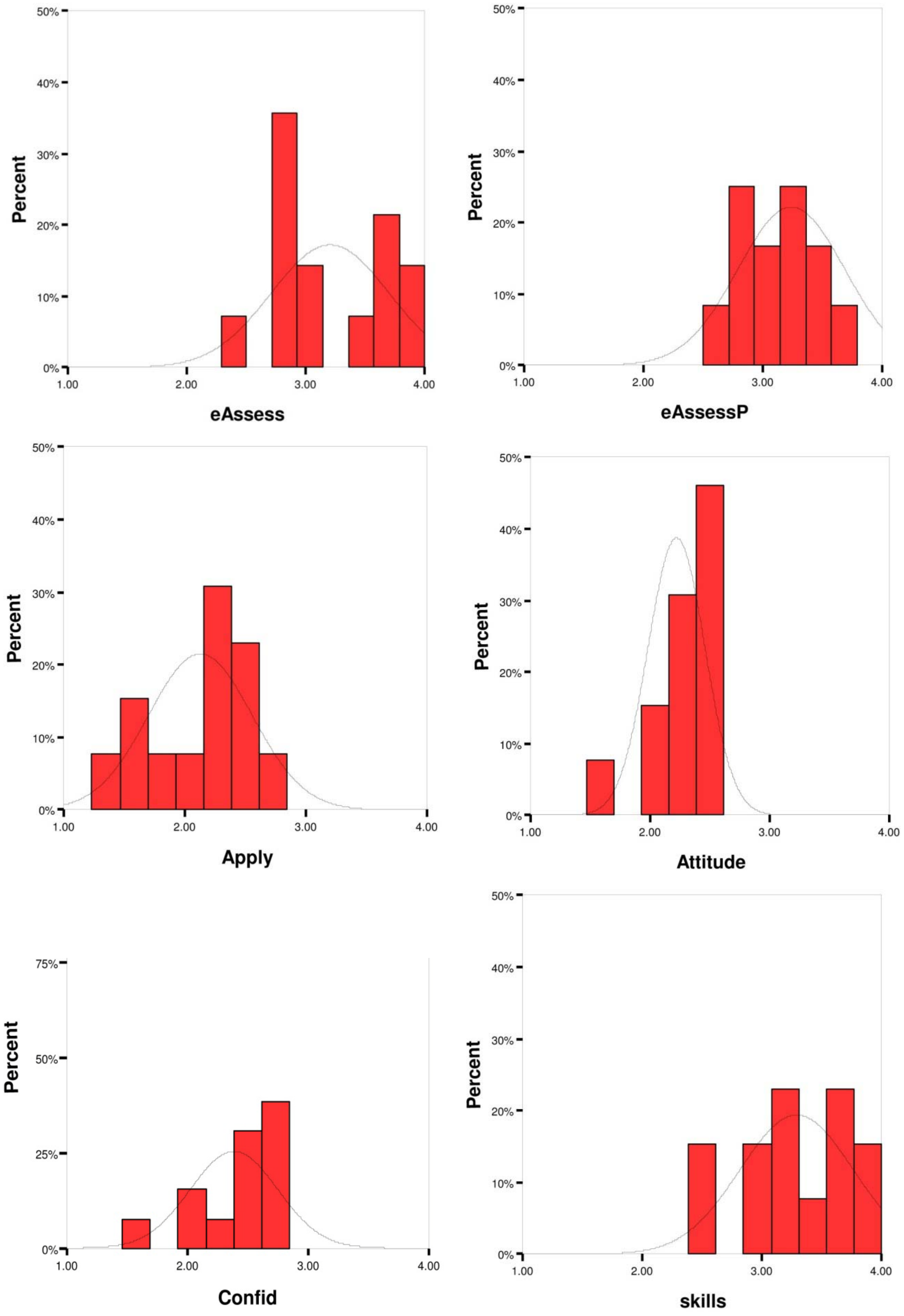


Figure 5.8 RA graphs for the distribution of scores for scales on the student questionnaire

Table 5.18

RA-Descriptions and Descriptive Statistics for Scales Based On Items from Student Questionnaire.

	N	Min	Max	Mean	SD	Description
eAssess	14	2.45	3.91	3.20	0.50	Perceived efficacy of computer use for the examination. Potential range 1 - 4
eAssessP	14	2.64	4.00	3.23	0.45	Perceived efficacy of computer use for the portfolio. Potential range 1 - 4.
Apply	13	1.40	2.8	2.12	0.43	Application of computer to various uses. Potential range between 1 and 3.
Attitude	13	1.60	2.40	2.22	0.24	Attitude towards using computers. Potential range between 1 and 3.
Confidence	13	1.50	2.67	2.38	0.36	Confidence in using computers. Potential range between 1 and 3.
Skills	13	2.45	3.91	3.29	0.48	Self-assessment of ICT skills. Potential range between 1 and 4.
SCUse	14	0.00	132.00	62.71	35.17	Estimate of time in minutes per day spent using computers at school.

Student Forum

Four students (3 male 1 female) agreed to stay behind immediately after the examination to be part of the student panel. The discussion focussed on the examination and not the portfolio and ran for 7m 30s.

What did you think of the task(s) you were asked to do?
After a bit of prompting students said they were familiar with these types of tasks; the exam tasks requiring the use of a variety of software applications to produce digital artefacts. With regard to the reflective questions on the portfolio one student said this was difficult as he hadn’t finished it (the portfolio) yet.

Did the computers help?
Given the timeframe they were able to produce work of good quality. They would have liked longer time for the practical and less for the reflective questions which “most people finished inside 20 minutes”.

How much different was this to how it used to be done?
“This was more practical and how we work in class time”. They said there was less stress doing it this way and that all round it was easier.

What, if anything, would you like changed in future?

“Nothing really…fine as it was” After prompting students said the task was reasonable for students at this level. They said again that Section A was too long. Wording? “OK. Pretty understandable what to do.”

Were there any technical problems with doing the activities?

There was widespread failure of sound recording hardware. Also, loading image files from samples to product took a long time for some students. This was apparently due to server lag / network congestion. One student lost the desktop altogether and couldn’t open any programs and had to move to a spare machine.

Any other thoughts or suggestions for developing the use of digital forms of assessment?

After prompting, all agreed that practical examinations would have student appeal and that working this way was “more modern than writing things”.

In summary, students were positive about the examination and considered it to be a fair, easier and less stressful method of assessment of AIT as well as being in line with what was done in class time. There were some adverse comments relating to the hardware and software deployed and the failure of the sound recording as well as suggestions for more flexibility in timing, with more for the practical and less for the reflective questions.

Email ‘Interview’ with Teacher

What did you think of the assessment tasks overall?

“The assessment tasks appropriate for my cohort of students. The students have developed additional skills in designing the artefacts and logos. I have used them for my skills assessment in this course”.

What did you think of the structure of the activities? “Timing is slightly a bit of a challenge, Students all work at different pace and sometimes they are absent from school, this caused some re organising for extension of time”.

What were the students' reactions to the activities?

Some students do not seem to accept the given time -line on completion of

> tasks… they needed more study skills and self-discipline. Most of the students enjoyed the tasks and seem to be able to work with little or no supervision. Where students finished them in a much quicker time, they needed to be given additional tasks.

What do you think of its potential? ?
“There is great potential for AIT and Engineering Studies. As we proceed further down this path, the process will improve. My students are telling me this should have happened earlier”.

What did you think of the quality of work produced by your students for these tasks?
“Some students have used *Photoshop* and came up with excellent examples of work”.

Were you surprised by the performance/attitude of any students?
“The students 14 of them sat the exam. They seem to accept that is pretty normal and would prefer that all their other exams are computerised”.
What was the general feedback from students?
“Happy with the exams and would like to see more of the same”.

Were there any technical problems with implementing the activities?
“Sound application and student familiarity of application software”.

Were there any other problems with implementing the activities?
“Sequence and timing of some tasks”

Any other thoughts or suggestions for developing the use of digital forms of assessment?
“Create a bootable USB memory stick and have all the required applications and programs for the exam so that students do not have to rely on a local network (LAN). In this way there is more control over the whole environment”.
In summary, the teacher was very positively disposed to both the examination and the portfolio and felt there was potential to develop computer-based assessment in other courses. The suggestion of a bootable memory device, holding resources and software

for the examination was an excellent idea and would certainly help to standardise the examination environment.

Results of Marking

The results for each student are shown in Table 5.19. The two external assessors marked five pieces of work for each student comprising the design document for the product, the product itself, two further digital artefacts, a theory section and a practical section of a three-hour examination. The marks for these were totalled with no allowance being made for missing work.

Table 5.19

Results for Case RA from Marking Portfolio and Examination (N=14).

	Ass Marking (Total %)			Assessors		Teacher(%)	Rank of	
St ID	Ass 1	Ass 2	Ave	Pfolio (70)	Exam (30)	Sem	Ass Ave*	Tch Sem
ra101	15	18	16.5	7.5	9.0	47	98.0	19.0
ra103	44	46	45.0	32.0	13.0	78	37.5	3.0
ra104	26	31	28.5	14.0	14.5	84	75.5	2.0
ra105	35	43	39.0	21.5	17.5	66	51.5	10.0
ra106	17	12	14.5	a	14.5	37	102.5	22.0
ra109	60	66	63.0	42.5	20.5	71	10.0	7.0
ra110	45	56	50.5	35.0	15.5	49	26.0	18.0
ra111	22	16	19.0	10.0	9.0	64	93.0	11.0
ra112	16	17	16.5	7.0	9.5	51	98.0	17.0
ra113	34	37	35.5	23.0	12.5	91	60.0	1.0
ra114	35	21	28.0	11.5	16.5	55	77.5	14.0
ra115	12	15	13.5	4.0	9.5	52	107.5	16.0
ra117	29	34	31.5	15.5	16.0	75	68.5	5.0
ra118	41	30	35.5	19.5	16.0	76	60.0	4.0
Mean	30.8	31.6	31.2	18.7	13.8	64.0	69.0	
SD	13.8	16.6	14.8	11.8	3.6	15.9	30.1	

**Ranking of external assessors is for all 115 students involved in the study.* a *Portfolio not submitted*

The teacher provided a semester score for each student and no separate portfolio or examination scores were supplied. The *rank-assessor average* is the position out of the whole sample of 115 students spread over seven schools. The *rank of teacher's semester (Tch Sem)* is the rank of the score awarded by the teacher out of 115. The mean of the average of the external assessors' marks (31.2%) was not statistically significantly different from the mean for all 115 students involved in the study (36.9% with SD of 17.2%). There was one student in this class ranked in the top 10% of students (*Rank Assess Ave*). The mean ranking of the external assessors was 69.0 which is above the mid-point of 58. Overall this class was diverse in ability but on average similar to the other classes.

A correlation analysis was done on the scores and on the rankings generated. The results are shown in the Tables 5.20 and 5.21.

Table 5.20

RA Correlation Coefficients from the Marking of Student Work (N=14).

	Assessor 1	Assessor 2	Assessor Average	Teacher Exam %
Assessor 1	1.00	0.90*	0.97*	0.46
Assessor 2		1.00	0.98**	0.44
Average			1.00	0.46
Teacher Exam %				1.00

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

Table 5.21

RA Correlation Coefficients from the Ranking of Student Work (N=14).

Rank of	Assessor 1	Assessor 2	Assessor Average	Teacher Exam %
Assessor 1	1.00	0.88*	0.96**	0.46
Assessor 2		1.00	0.98**	0.47
Average			1.00	0.47
Teacher Exam %				1.00

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

Correlation between the external markers was strong with a coefficient of 0.90 ($p<0.01$) for the marking and 0.88 ($p<0.01$) for the ranking of students. When compared with the average of the two assessors, the correlations of the teacher's marks for the semester are weak and not significant.

Conclusions

Conclusions pertaining to this case study are discussed separately for the portfolio and the examination.

Portfolio

From the survey it was clear that students were positive towards computer-based portfolios and clearly welcomed assessment of their practical capabilities. They indicated some familiarity with this type of assessment. There was a wide variation in the quantity and standard of work submitted for the portfolio. Often, promising products were let down by broken or relative hyperlinks and missing pictures which probably all worked well on the student's computer but failed when zipped and submitted for marking. Both the students and the teacher were comfortable with this form of assessment. Students indicated that they could fully demonstrate their understanding, skills and knowledge through performance of on real world activities of this type. The portfolio tasks mirrored typical class activity and were observed to be easily manageable.

Examination

A number of technical difficulties and delays were experienced connected to the fact that application software was run from a central server rather than on the local machine. On three occasions, student computers stopped responding and had to be restarted. Fortunately there were additional computers available and students were able to quickly relocate to these. The audio recording was also problematic. Sound card drivers were disabled here by group policy. However, a visit the day before the examination met with assurances that sound cards had been enabled.

The student forum agreed that more time for the practical would have been beneficial perhaps reflecting the delays experienced. The students responded well to the style of examination and appeared comfortable with the practical component. Many students didn't seem to appreciate the difference between a logo and a banner or poster and there was much reliance on the supplied images. Numerical data was well handled using a spreadsheets and students were clearly familiar with graphing. Although agreeing that the theory section was made easier by being done on a computer, many students found the reflective questions difficult to understand and complete. Many appeared to misunderstand the intention of the questions and to which part of the technology process they referred. Many answers were off topic.

There was a high level of agreement between the two external assessors but their average bore little relation to the teacher's assessment both numerically and ranked. The teacher made the suggestion of including not only the working files on the mass storage device but also a range of application software. With 4 GB of memory, it would be possible to include full office and multimedia applications running locally. This might present a more *level playing field* for all candidates.

Case Study WA: Private School

The WA case study involved one teacher and a class of 14 year 11 boys. The class was conducted in a computer laboratory in a modern purpose built centre as pictured in Figure 5.9.

Figure 5.9 WA computer laboratory

The context for the class was *Business Information Technology* and the portfolio product was modified by the teacher to the development of advertising for a real client, a hair dressing studio close to the school. The manager of the studio had visited the school and met with the class to outline her requirements and specifications. Previously, students had developed a website for the same client incorporating the specific colours, themes, styles and images of the business. The current requirement was for some form of advertising for which the teacher developed a design brief for a thirty second cinema commercial.

Implementation, Technologies and Issues Arising

The researcher either met or communicated using phone and email with the teacher before the students became involved. This was to discuss the process with the teacher and to test the technologies, in particular the use of computers in the school's laboratory with sound recording and a USB flash drive. Computers were up to date but there were issues concerning the Standard Operating Environment and rights on the network which were highlighted by the teacher.

Data Collected

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 assessment task.

Observations of the Classes

The class was visited on four occasions to observe students completing the assessment tasks, or to collect qualitative data.

Visit 1: Product Development (14/08/08, 9:30-10:40)

There were 14 boys present in the computer laboratory working for about one hour on the first component of the portfolio, the digital product. Work had commenced the previous week. The students were divided into four groups on the basis of experience and ability, with one group of *better* students working on a project that involved development of an online database for which the students needed to learn some *PHP* and *SQL*. The other three groups were each working on the designs for a hair salon owner (a real local business). The owner had supplied a design brief that included colour preferences, photographs, trophies and information required. One of the groups was of *less able* students so the teacher needed to provide them with help.
Students worked independently on individual tasks for their group with decisions already made about group roles. The next day the teacher intended to show the salon owner the prospective designs that the three groups had developed. Therefore at this stage each group had concept design files (e.g. logos, banners, layouts). Students used *Adobe Fireworks* for graphics and *Dreamweaver* for web pages. For technical reasons, there was no access to students’ server space so USB drives were used to transfer and save files.

One student used *Inspiration* to create a concept map. All had access to a common set of photos from the hair salon *Peta Charles* in a shared area on the server. Most students were working on graphics and a few on web pages but about half got little done during session.

Visit 2: Development of Process Document (06/11/08, 9:30-10:40)

There were 14 boys present in the computer laboratory. The teacher gave out the *Design Document Template* and the researcher explained again the requirements of the project. Students listened intently and appeared to understand the purpose of the research and their part in it.

Students were working on a cinema advertisement for a real client, a local hair studio. The client had visited the class to discuss her requirements and provided guidelines on colour schemes as well as digital photographs of some products and services. Students were developing the advertisement in two main formats: *Flash* and *PowerPoint*. Students using *Flash* said they had learned mainly from each other and although this was claimed to be their first exposure, the work appeared very sophisticated, with multiple layers and motion *tweening* yielding some very professional looking work. *PowerPoint* work also looked good. Though simpler to use, this product is quite sophisticated and by using media imports and the many effects available together with rehearsed timings of the slide show, a perfectly presentable result was possible. One student was using *Moviemaker* which again was an easy to use and very powerful tool. All products were nearing completion and due in that week. The teacher said that the intention would then be to have the students complete the design document and the practice exam before the final exam. The teacher and researcher discussed the latter and it transpired that some sound card drivers were missing. A sample headset was left for testing.

Visit 3: Examination (27/11/08, 1:30-4:30)

Fourteen students sat the examination in the computer room. The sound test and software tests prior to commencement did not go well. Many machines would not record sound or had missing software. The end result was that the group of 14 was split into two with seven students moving to an adjacent laboratory where sufficient working machines were available. This group ran the test again and their examination started ten minutes behind the other group. Sound recording still appeared to be problematic. The exam began with the 2 hour practical component. After 20 minutes, more technical problems came to light as *Adobe Fireworks* failed to open for student WA105. He moved to another machine but this caused a delay for him and a note was made to add time on. As well as *Adobe Fireworks s*tudents also used *Flash* for the logo.

Students asked an unusual number of questions mainly about what was required and how to proceed. None appeared to have *Publisher* open or used a template for a brochure although they were available. Also some had difficulty with the graphing and with what a tri-fold brochure actually was. It was noticeable that a few students were filling in the designs retrospectively. Some even used the product created in the supplied design template.

When the time came to record the audio reflection only a few students were able to do so. The teacher and researcher decided to use the teacher's laptop to record the audio reflections one at a time and the teacher set this up outside the rooms. Students took their USB flash memory with them. Eventually they all got through the recording whilst section B the reflective questions continued. Nobody looked at the portfolio product when answering Section B which was strange as the questions referred to it.

Visit 4: Survey and Forum (27/11/2008 4:30-5:00)

On completion of the exam, students completed the questionnaire. A group of four students agreed to take part in a student forum. They were presented with the same set of questions as other cases but follow up questions differed depending on responses.

Survey of Students

The survey was completed by all 14 students immediately on completion of the performance examination. The minimum, maximum, mean and population mean (mean for all 115 students) were calculated for each closed response item using SPSS (refer to Appendix K).

Items Concerning the Portfolio

Students reported varying degrees of previous experience of completing computer-based portfolios. Though the mean for this item was 2.2 on a scale of 1 (Lots of experience) to 4 (No experience) five students said they had little or no previous experience. The population mean for this item was 2.5 indicating that this group was slightly more experienced than the norm. Students indicated that they would require some *time to get used to it* (item P1(b)) with a mean of 2.5 on the same scale. Strangely one student reported lots of experience but also lots of time to become used to completing a digital portfolio. Students also agreed or strongly agreed with the statements in survey items P2(a)-P2(f) which asserted that the computer was *quick, easy* and *good* for developing their *ideas* and completing the portfolio *product* with means for these items being in line with the population, ranging from 1.6 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students' preference for computer-based portfolios was also very strong with responses to items such as *it was better doing the portfolio on a computer than on paper* and *overall the computer is good for portfolios* producing means of 1.3 and 1.4 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to population means of 1.6 for the same items, indicating that this group was even more positive than the norm. Only in response to item P2(h), *the steps of the portfolio helped me develop my ideas* was any disagreement apparent; three students disagreed with the statement and the mean for this item was 1.9 in line with the population mean of 2.0.

Items Concerning the Examination

Twelve of the 14 students indicated that they had little or no previous experience of completing an examination on computer. The mean for this item was 3.2 on a scale of 1 (lots of experience) to 4 (no experience) even though the teacher had been supplied with a practice examination of similar scope and difficulty. Responses to the item E1(b) *how much more time would you need to get used to it?* were varied with a mean of 2.6 on a scale of 1 (Lots of time) to 4 (No time). Students were positive about using the computer for the examination, agreeing or strongly agreeing with statements E2(a) and E2(c)-Eq2(i) suggesting it made things *quick* and *easy* and was *good* for creating the logo, graphs and brochure. Means for these items ranged from 1.4 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree) largely in line with population means. Only statement E2(b) *it was easy to use the computer in the exam to develop ideas* produced any noticeable dissent with five students disagreeing and one strongly disagreeing with this assertion. The item, *overall I was able to show what I could do in the exam,* produced a mean of 1.9 in line with the population mean. Students' preference for a computer-based examination was very strongly supported with the statement *it was better doing the exam on computer than on paper* producing a mean of 1.3 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to 1.7 for the population as a whole.

Items Concerning Computer Use

These students were regular and extensive users of new technologies with 13 out of 14 using a computer at home on a daily basis. All students had home broadband internet as well as a full range of home entertainment and communication systems with all but one owning a mobile ‘phone and all but two an MP3 player. Attitude to computers was definitely positive with widespread ease of use and facility reported. Students self-assessment of their proficiency with application software such as word processing and spreadsheets was also high with means ranging between 3.4 and 3.7 on a scale of 1 (low ability) to 4 (high ability); the only exceptions being databases and digital video editing where means were both 2.7 in line with population means of 2.6 and 2.9 respectively.

Open-ended Items

A summary of responses to the four open-ended questions on the student questionnaire is shown in Table 5.22.

Table 5.22

WA Summary of Students' Responses to Four Open Ended Questions on Survey Questionnaire

The two best things …	The two worst things …
Portfolio Eight of the 28 possible responses were left blank. *Easy, Easier, Simple, Neat* featured in six responses, e.g. “Easy to show work” and “Simple to develop”. *Better* was alluded to in responses such as “Computers are better at doing stuff more professionally” and “Show cases your skills” and “You actually learn things”.	Portfolio 14 of 28 potential responses were left blank. Here was no dominant theme. Some sample responses were “Computers can freeze easily” “Not enough time in the course “ “Frustrating when design comes out wrong, especially when using advanced programs” and “Can get sick of it (get bored fast)”.
Examination Easy (7 students used the word easy or easier) e.g. “Easy to produce final designs” and “Easier to type than write” Appropriate (3 students made reference to the suitability of the examination) e.g. “Using the computers is very practical and fitting to the subject” and “The design is in front of you and you can actually show your skill” Practical (3 students made reference to the practical nature of the examination) e.g. “Shows what we can actually do in AIT there isn't just theory”	Examination Technical problems (6 students made reference to technical problems with hardware and software) e.g. “Computers wont set up properly at first, wasted some time” and “The programs wouldn't work” Time (3 students made reference to a lack of time) e.g. “Not enough time” and “lack of time”.

Students considered that using computers for the portfolio and the examination made it easier both for designing and developing their product. They welcomed the inclusion of the practical examination component which they felt allowed them to demonstrate skills which would otherwise go unexamined. The main worst things were concerns about the reliability of the network and the inconsistent nature of the configuration of the computer that they had been allocated. Several students found necessary applications to be missing and had to move to an adjacent room during the examination and this is reflected in their comments.

Questionnaire Scales

Seven scales were derived from combining selected items from the questionnaire. An explanation of the scales is set out in Table 3.4 in Chapter Three. The results are shown in Table 5.23 and graphically in Figure 5.10

Table 5.23

WA-Descriptions and Descriptive Statistics for Scales Based on Items from Student Questionnaire

	N	Min	Max	Mean	SD	Description
eAssess	14	2.73	3.82	3.27	0.34	Perceived efficacy of computer use for the examination.. Potential range between 1 and 4
eAssessP	14	2.73	4.00	3.34	0.40	Perceived efficacy of computer use for the portfolio. Potential range between 1 and 4.
Apply	14	1.67	3.00	2.54	0.33	Application of computer to various uses. Potential range between 1 and 3.
Attitude	14	2.00	3.00	2.65	0.21	Attitude towards using computers. Potential range between 1 and 3.
Confidence	14	1.50	3.00	2.67	0.38	Confidence in using. Potential range between 1 and 3.
Skills	14	2.27	3.91	3.27	0.48	Self-assessment of ICT skills. Potential range between 1 and 4.
SCUse	14	48.0	180.0	100.35	44.60	Estimate of time in minutes per day spent using computers at school.

Students indicated that the examination and portfolio were both appropriate assessment instruments with relatively strong computer efficacy with the means for both *eAssessP* and *eAssess* of 3.3 being approximately 2 and 2.5 standard deviations above the mid-points.

Attitude and *Confidence* scales were also positively skewed with means for both of approximately 2.7 being three and two standard deviations above the midpoints respectively. Students were confident in using ICT (mean=2.7, midpoint=2.0, SD =0.4) and had a high (self-assessed) level of ICT skills across a range of applications (mean=3.3, midpoint =2.5, SD =0.5). On average they indicated using ICT for more than 1 ½ hours per day at school.

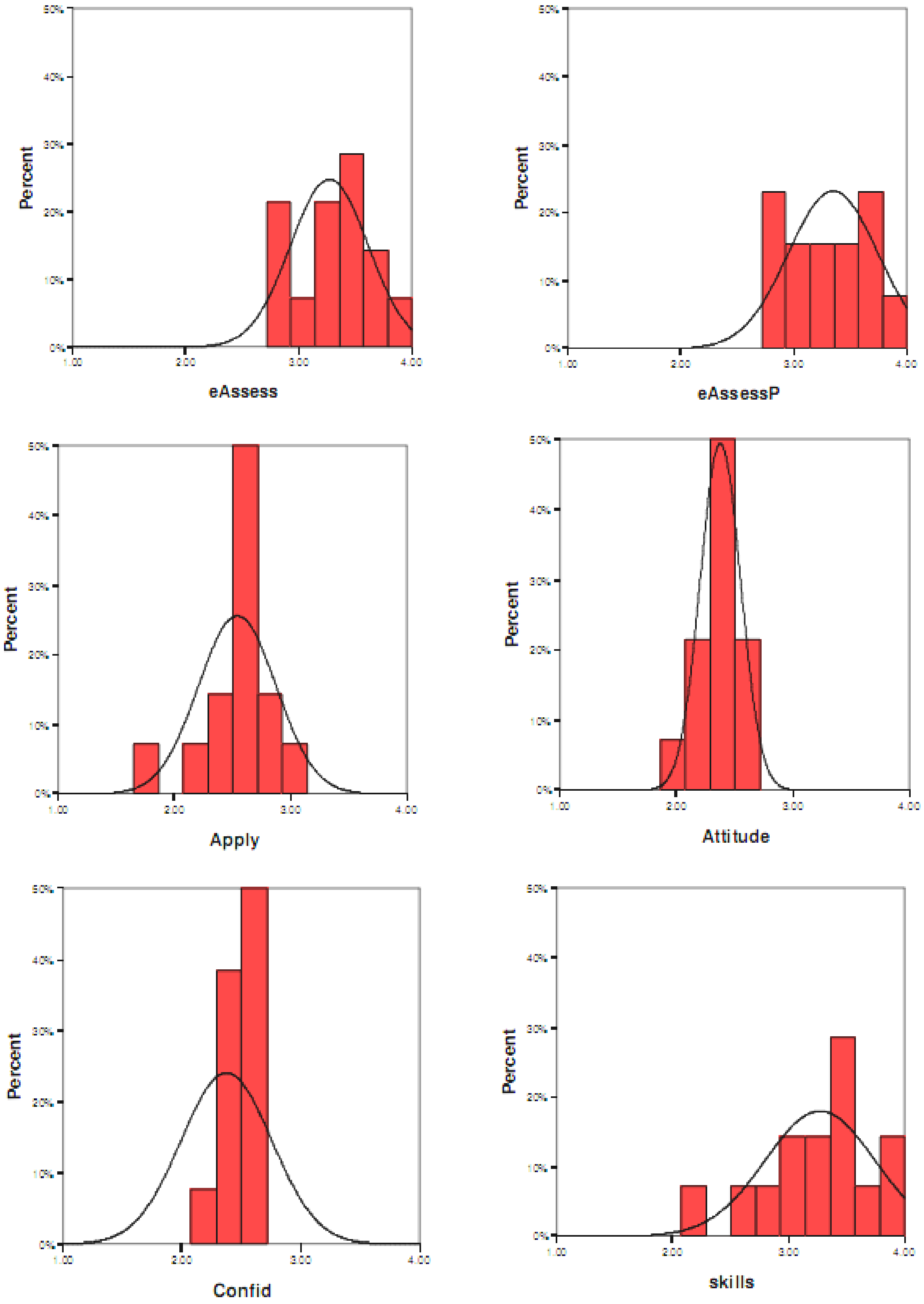


Figure 5.10 WA graphs for the distribution of scores for scales on the student questionnaire

Student Forum

A group of four students was interviewed immediately after the examination with discussion focussed on the examination and not the portfolio. The discussion ran for 6m 20s.

What did you think of the task(s) you were asked to do?

“pretty easy…I thought the last section [reflective questions] was a bit confusing sometimes”. “The last part…about the wording” All concurred that the language was a bit too advanced and this made it difficult. The general feeling here was in the negative with the reason being attributed to technical difficulties with software. “It could have had a bit more time” “more time for planning and producing”. “I spent too much time on the design” “I think at the end…if you have any time left over you should be free to go back”.

Did the computers help?

There was a general agreement that the computers helped. “It made the design part much easier for me as I can’t draw well” “Definitely made it easier [to complete the reflective questions]”.

How much different was this to how it used to be done?

“The [reflective] questions are the same it’s just the way you do them.” “There’s more variation in what you can do”. When prompted about the structure of this examination, (a major practical together with a theory component), all students strongly agreed this was better.

What, if anything, would you like changed in future?

“Time allocations [more time and unrestricted use of time]”. “Talking into the computer ... that didn’t really work”. “Speaking out an answer could be a lot quicker”.

Were there any technical problems with doing the activities?

Applications in this school ran from a server and this presented some problems. “Most of the programs at the start wouldn’t even download”. “Needs to be all set up beforehand to make sure”.

Were there any other problems with the activities?
"The exam…was mostly just design with no emphasis on other subjects [areas of the course] like databases and spreadsheets". "We've done quite a bit of the technical stuff and not much on design".

In summary, students were positive about the practical aspects of the examination considering them to be a straightforward, fair and comprehensive method of assessment of AIT. There were some criticisms of the reflective questions which were perceived as confusing and repetitive. Students did not support the audio section suggested a short interview in its place and were worried about systems failure. They also would have preferred to manage their own time and have been able to go back to complete unfinished work.

Email 'Interview' with Teacher

What did you think of the assessment tasks overall?
"I thought the tasks were excellent, easy enough for less able students to achieve an adequate result but open ended enough for the good students to excel".

What did you think of the structure of the activities?

> It seemed inappropriate to suddenly interrupt them and tell them they now had to do something else, like plan or record a response. It seemed unfair not to let them go back and add to work done in Part A after the time for that portion of the exam had expired. Design is an iterative process, and the ability to review and improve your work is a work habit to encourage. Time allocated was fine, but I had a major problem with the marks allocation. 30 minutes and 45 minutes to create a logo and a brochure, but the marks allocated were paltry. 10 plus 5 minutes to plan and record a response and this earned as many marks as the 75 minutes allocated to creating product. Making the sum of marks for the exam equal 30 was artificial. Not enough scope to differentiate between levels of performance-I thought the idea was to honour skill but there seemed to be an absurd emphasis on writing about what you did, and the display of skill was undervalued.

What were the students' reactions to the activities?

“There were unhappy with the marks distribution. They said the exam was easy, but most performed surprisingly poorly”.

What do you think of its potential?

> Limited potential. We encountered a ridiculous number of technical problems, I expected a few and had a contingency plan, but the number of machines that would not record sound or run software was a shock. I checked all machines 4 weeks before the exam and 20/22 were fine. 4 weeks later it was a shambles. There is way too much potential for seeing what other people are doing and get ideas from others. Any student with the intention of cheating could have done so too easily.

What did you think of the quality of work produced by your students for these tasks?

> I thought the quality of work was generally poor. When the students were reflecting on projects done earlier in the year they failed to mention many of the steps they took in developing their solutions. They did not do themselves justice at all.

Were you surprised by the performance/attitude of any students?
“Despite the difficulties experienced their attitude was positive. I strongly believed they were getting an excellent opportunity to show what they could do, and sold the idea pretty well”.

What was the general feedback from students? [No response]

Were there any technical problems with implementing the activities?
“Many of the USB files were corrupt. One student had to try 3 thumb drives before he got all the files onto his computer. Heaps more. See above”.

Were there any other problems with implementing the activities?
[No response]

Any other thoughts or suggestions for developing the use of digital forms of assessment?

> I marked the exams as per the marking key, though I did not feel that the marking key was fair or relevant. Then I studied the distribution of marks, and made a quick pair wise comparison based on quality of design of the logo and brochure. I decided the original marking scheme was broken. I remarked all exams again allocating 1 mark for every minute suggested -except for the last question -60 minutes allocated but I marked out of 30. This yielded a distribution of marks much closer to what I expected, and, except for two students whose marks were better under the original marking scheme, it was these marks I used to calculate grades for the course.

In summary, the teacher's comments appeared to be applied to the examination only describing the tasks themselves as excellent. The implementation and marking elicited several qualifying statements. Understandably, given the number of technical issues encountered, sentiment was not as positive as for other cases. The teacher had concerns about the opportunity for students to copy ideas from one another and was not at all happy with the marking key which was perceived as being out of line with the tasks themselves by placing too much emphasis on the design and not enough on the creation of the product. In principle the teacher was supportive of the intention and type of assessment with the details of managing, and timing needing attention.

Results of Marking

The two external assessors marked the theory and practical sections of the three-hour examination. The marks for these were totalled and the results of the analytical marking for each student are shown in Table 5.24. The mean of the average of the external assessors' marks *(Ass Av*) on the exam component was 12.6 (42%) slightly higher than the mean for all students of 10.2 (34%).

A correlation analysis was done on the scores and on the rankings generated by the assessors. The results are shown in Table 5.25 and Table 5.26. Correlation between the external markers was moderate with a coefficient of 0.66 ($p<0.05$) for marking and 0.57 ($p<0.05$) for ranking. The marks awarded to this sample were in general very low due mainly to the fact that the portfolio was not submitted; hence only the practical and theory sections of the examination were assessed, the other components being scored as zero. No teacher marks were received either for the examination or for the portfolio even though when interviewed, the teacher indicated that these had been marked. As no portfolio work was submitted and no teacher marks received, no further analysis could be undertaken.

Table 5.24

Results for Case WA from Marking Exam (N=14)

Student	Assessor marking (Exam only)			Rank
	Ass 1 (/30)	Ass 1 (/30)	Ass Av (/30)	Ass Ave*
wa101	15	14	14.5	62.0
wa102	16	11	13.5	73.0
wa103	14	13	13.5	73.0
wa104	10	11	10.5	95.0
wa105	14	15	14.5	62.0
wa106	12	15	13.5	73.0
wa107	16	17	16.5	42.0
wa108	9	4	6.5	111.0
wa109	10	12	11.0	88.5
wa111	11	11	11.0	88.5
wa112	12	13	12.5	81.5
wa113	7	9	8.0	107.5
wa114	18	14	16.0	48.0
wa115	14	15	14.5	62.0
Mean	12.7	12.4	12.6	76.2
SD	3.1	3.2	2.9	20.5

*Ranking of external assessors is for all 115 students involved in the study.

Table 5.25

WA Correlation Coefficients from the Marking of Student Work (N=14)

	Assessor 1	Assessor 2	Assessor Average
Assessor 1	1.00	0.66*	0.91**
Assessor 2		1.00	0.92**
Assessor Average			1.00

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

Table 5.26

WA Correlation Coefficients from the Ranking of Student Work (N=14)

Rank of	Assessor 1	Assessor 2	Assessor Average
Assessor 1	1.00	0.57*	0.86**
Assessor 2		1.00	0.87**
Assessor Average			1.00

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

Conclusions

Conclusions pertaining to this case study are discussed separately for the portfolio and the examination.

Portfolio

Students were very positive about assessment by digital portfolio in the survey claiming it was quick, easy and good for developing ideas and preferable to paper based alternatives. Although the teacher was very comfortable with this form of assessment, students reported varying degrees of familiarity but felt this was a fair method of assessment allowing them to do full justice to their skills. There was a wide variation in the quantity and standard of work observed. The product, multimedia advertising, was attempted using a variety of application software by students of widely different ability levels. No working products were submitted although partially developed products were observed during visits.

Examination

The students responded well to the style of examination and appeared comfortable with the practical component. Many students didn’t seem to appreciate the difference between a logo and a banner or poster and there was much reliance on the supplied images. Numerical data was generally well handled but a few students didn’t appear to know how to use a spreadsheet for graphs and charts.

A number of technical difficulties and delays were experienced. For example about half the students found that the required application software was missing from their computers and had to be moved to an adjacent computer laboratory. This took ten minutes and the two groups had to have timings adjusted accordingly. The audio recording was also problematic. On this *Novell* network, student rights to sound cards had been disabled. The teacher had requested in advance access to audio and was assured this had been done. Only when sound tests were done prior to commencement of the examination was it discovered that audio recording was not possible. Students completed this section one at a time outside the computer laboratory by recording on the teacher’s laptop.

Although agreeing that the theory section was made easier by being done on a computer, many students found the reflective questions difficult to understand and complete. Many appeared to misunderstand the intention of the questions and to which part of the technology process they referred. There were several requests for clarification over wording, surprising in an examination, which were not entertained. Many answers were off topic

This case was problematic in implementation, with manageability and technical difficulties; in fact a *worse* case scenario would be hard to contemplate. The teacher was highly experienced, capable and enthusiastic, at least at first, and appeared to be frustrated and let down by network support staff. Nevertheless, the examination was eventually completed successfully and student work samples were collected for assessment. Though students were inconvenienced and upset by technical failures they still claimed to have demonstrated their best work under the time constraints allowed.

Case Study XA: Private School

The XA case study involved one teacher and the 29 year 11 students from two mixed 11/12 classes in a Business Information Technology context. The classes were conducted in one of the two adjacent computer laboratories pictured in Figure 5.11.

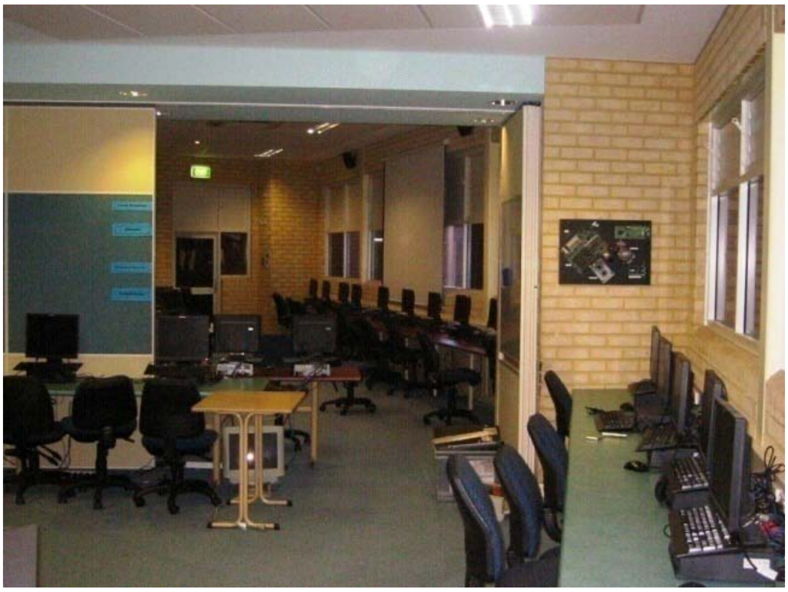


Figure 5.11 XA computer laboratory.

The dividing partition allowed the two labs to be invigilated as a single room during the examination. The teacher followed the design brief supplied, changing the portfolio product to a short animated feature on a health and safety issue relevant to Year 11 students. Further, students were allowed to work in small groups on their portfolio product and were jointly assessed by the teacher. Permission had to be sought from school administration to allow the performance examination to replace the usual written examination. The examination was scheduled and invigilated during the end of semester examination period and counted towards the student's final semester mark.

Implementation, Technologies and Issues Arising

The researcher was a colleague of the teacher and meetings took place frequently to discuss progress and to test the technologies, in particular the use of computers in the school's laboratory with sound recording and a USB flash drive There was a marked performance difference between the computers in the two labs due to a difference of two years in the age of the machines. Though both ran the same Standard Operating Environment and had the same amount of memory, the more modern machines were fitted with a faster and dual core processor. This assisted multitasking application and processing large image files.

Data Collected

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 assessment tasks.

Observations of the Classes

This case was slightly atypical in that both of the two classes were observed working on one aspect of the portfolio. For the examination, both classes were merged and so observations and composition of the student panel derived from the combined group.

Visit 1 Product Development (22/08/08 8:45-9:45)

The group observed in this initial visit was one of two classes of AIT Unit 2B made up of mixed year groups. Nineteen students were present, with two absent; 12 of these were the Year 11 students taking part in the study and seven were Year 12 students. Work had commenced the previous week with students working in groups of two or three on the portfolio product to a brief developed by the teacher. The task was to create a safety animation targeted at teenagers using some form of animation. Most had selected *stop-motion claymation* or *stop-motion* using physical toy models. The action was captured with webcams mounted on tripods, with characters in front of cardboard backdrops, using *Image Studio* to do image sequences and *Adobe Premiere* to join sequences and add audio. Other students were observed to be using *Flash* or *CorelRAVE* to create their animations. *Adobe Photoshop* was used to create backgrounds both printed and digital.

All seemed very engaged, with most working on creating the animation with a few completing the storyboard to hand in to the teacher. Some storyboards were done digitally using *MS Word* or *PowerPoint*, but most were done on paper with pencil drawings. The Teacher demonstrated the use of *Adobe Premiere* to combine videos and add an audio track. For the second half of the lesson a radio played background music. Most students appeared to be following the script of their storyboard and had the paper in front of them or the *MS Word* document open. One boy printed a coloured net for a 3D building and a ‘hillside’ scenery backdrop.

Visit 2 Development of Process Document (15/09/08,11:30 - 12:30

The second class of 23 students (including eight year 12 students) was chosen for the visit to observe work on the *Design Process Document* in the computer laboratory. The class had recently started work on this and the lesson began with a re-explanation by the teacher of the requirements of the document. The template for this was displayed on a data projector. Students seemed clear on the requirements of the process document and began or continued working. Some made reference to their designs for the animation and the research notes they had made. A few students were still working on the animation itself and it appeared that many of these were not yet complete; though the shooting of scenes was done, the final edits still needed work. The class were fully engaged for the whole hour.

Visit 3 Examination (20/11/08 1:00- 4:30

Due to a mix up with the class listings, three additional students presented for the examination so extra USB mass storage and headphones were required and these were delivered to the school in time for the start. Twenty-nine students were present and these were dispersed through the two adjacent computer labs with the partition wall partially open. Students were well spread out with most separated by an unoccupied workstation. There were two initial logon problems prior to reading time and these were resolved by moving students to spare workstations. This made the distribution of students somewhat uneven and in places spacing became a little cramped. Students' screens were turned to make adjacent viewing angles as wide as possible but the potential for overlooking could not be eliminated with any certainty.

Prior to commencement, all headphones and USB flash drives were successfully tested by students. Reading time commenced, with students permitted to browse the USB drive contents. The examination began with the obligatory 15 minutes planning section. All students initially did their planning on paper only one or two were observed to even open the templates. A few looked at the photograph folder and data file. Two boys started using the *PowerPoint* design template. One girl folded paper to check the layout of a tri-fold brochure. After the planning section, one girl took advantage of the active internet connection and opened the online *Title Producer* to do a fancy brochure title. For the logo many used *Photoshop*, some used *Adobe Fireworks* or *PowerPoint*. Almost all were observed to use a *Publisher* template for the brochure. At various stages, four of the older computers froze when opening *Photoshop* and had to be restarted and though no work lost this was understandably frustrating for the students. These machines were all in the same room and being older and inferior in system resources were observably slower in the processing of image files, in effect disadvantaging the students allocated to them. One boy went to *Wikipedia* and *Google Images* prompting students to be reminded with a general announcement that they could only use the supplied photos and graphics and were not to download any other material even though internet access was available throughout the examination. Most students used *Adobe Fireworks* or *Photoshop* to begin working on the logo. Some used *Paint*. Many took cues from photographs and used or adapted bits of these to make a logo. The turtle photograph was quite prominent in several logos.

Graphing with *MS Excel* seemed to be well understood by the group as a whole though no student was observed to attempt a combined rainfall (columns) and temperature (line) graph. For the brochure, almost all selected a template from *MS Publisher* and dropped in images, many adding their own text and slogans. Some good work appeared to be done.

Visit 4 Survey and Student Forum (20/11/08 11:30-12:00

On completion of the examination, 29 students completed the survey questionnaire. A group of five students (2 girls) agreed to stay behind and take part in a student forum. They were presented with the same set of questions as other cases but follow up questions differed depending on responses.

Survey of Students

Twenty-nine students completed the survey questionnaire consisting of 70 closed response items and four open-response items. The minimum, maximum, mean and population mean (mean for all 115 students across the seven cases) were calculated for each closed response item using SPSS (refer to Appendix K).

Items Concerning the Portfolio

Students reported varying degrees of previous experience in completing computer-based portfolios (with responses evenly spread across the scale of 1 (Lots of experience) to 4 (No experience) and a mean of 2.2. The population mean for this item was 2.5 indicating that this group was slightly more experienced than the norm though eight students said they had no previous experience. Students indicated that they would require some *time to get used to it* (item P1(b)) with a mean of 2.8 on a scale of 1 (Lots of time) to 4 (No time). Students also agreed or strongly agreed with the statements in survey items P2(a) to P2(i) which asserted that the computer was *quick*, *easy* and *good* for *developing* their *ideas* and *completing* the *portfolio product*, with means for these items being in line with the population, ranging from 1.6 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students' preference for computer-based portfolios was strong with responses to items such as *it was better doing the portfolio on a computer than on paper* and *overall the computer is good for portfolios* producing means of 1.4 and 1.8 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to population means of 1.6 for the same items. Only in response to item P2(g) *It was easy to follow the steps to create the portfolio* was any disagreement apparent; seven students disagreed with the statement (two strongly) and the mean for this item was 2.2 compared to the population mean of 2.0.

Items Concerning the Examination

Students reported varying degrees of previous experience in completing computer-based examinations with responses evenly spread across the scale of 1 (Lots of experience) to

4 (No experience) with a mean of 2.6 compared to a population mean of 3.2. Sixteen students reported lots or some experience. The teacher had been supplied with a practice examination of similar scope and difficulty but whether or not this was completed by students was unknown. Responses to the item E1(b) *how much more time would you need to get used to it?* produced a mean of 3.0 on a scale of 1 (Lots of time) to 4 (No time) indicating that students were in general comfortable with the examination; only six students responded that they would require some or lots of time.

Students were positive about using the computer for the examination, agreeing or strongly agreeing with statements E2(a) and E2(c) to E2(i) suggesting it made things *quick* and *easy* and was *good* for creating the logo, graphs and brochure and *reflecting on ideas* means for these items ranged from 1.3 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree) largely in line with population means. Only statement E2b*, it was easy to use the computer in the exam to develop ideas,* produced any noticeable opposition with eight students disagreeing with this assertion (mean=2.1, population mean=2.1).The item, *overall I was able to show what I could do in the exam,* produced a mean of 2.0 in line with the population mean of 1.9 only three students disagreeing one strongly. Students' preference for a computer-based examination was strongly supported with the statement *it was better doing the exam on computer than on paper* producing a mean of 1.6 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to 1.7 for the population as a whole. Only five out of 29 students disagreed, one strongly.

Items Concerning Computer Use

Students indicated regular and extensive use of information and communications technologies with all 29 reporting computer use at home on a daily basis. All students had home broadband internet as well as a full range of home entertainment and communication systems with all owning a mobile 'phone. Attitude towards and confidence with computers was definitely positive with all but 2 students asserting *I'm good at using computers*. Students' self-assessment of their proficiency with application software was varied. Whilst high ability was reported with word processing, slideshows, email, file management, internet research, digital photography and image editing (means 3.4 to 3.9) students were less confident with spreadsheets, databases and web authoring (means 2.6 to 2.9) on a scale of 1 (low ability) to 4 (high ability).

Open-ended Items

Responses to these items were tabulated to assist in drawing out themes and Table 5.27 shows a summary of the main responses.

Table 5.27

XA Summary of Students' Responses to Four Open Ended Questions on Survey Questionnaire

The two best things …	The two worst things …
Portfolio	Portfolio
The use of words such as *simple easy* and *quick* predominated. For example "It's easier and quicker" and "Easier to collate data/previous work" and "It's easy to get used to". Other themes were enjoyment, for example "It was fun" and "It was an enjoyable assignment" and reference to the type of product, in this case a movie, for example, " making a movie" and "filming"	A theme was poor hardware often combined with fear of loss of work or computer crashes. For example, "The computers are prone to freezing resulting in loss of work" and "Your portfolio can be lost easily". And "The file corrupted and I had to start again". Here was one reference to group work "My partner was not always there to help me" and two to the repetitive nature of the shot by shot movie making task, for example "Once your used to the programs it gets repetitive" and "Boring to do". Almost half (27/58) responses were blank.
Examination	Examination
Easy, Quick and Comfortable (20 students made reference to one or more) e.g. "So much easier" and "Quick and efficient in designing logos" and "Comfortable environment"	Hardware (17 students referred to delays, freezes, hangs, and a general dissatisfaction with the capability of the hardware) e.g. "Slow when lots of applications are opened" and "Sometimes the computer will go slow and freeze wasting time and sometimes causing work to be lost"
Enabling (5 students made reference to the suitability of the task) e.g. "We were able to show our ability using computers" and "Able to develop an actual product".	Noise (6 students mentioned the irritation of hearing key tapping by neighbouring students). e.g. "I didn't like the tapping on the keys from other students in the exam"

Generally students considered that using computers made the portfolio and the examination quicker, simpler, easier and more relaxed. The computer laboratory was seen as a preferable and more suitable environment within which they could demonstrate their skills and their ideas. The main worst things were a concern that the computers were not capable of running the application software at the required speed and might crash. Two students had to relocate during the examination due to missing application software.

Questionnaire Scales

Seven scales were derived from combining selected items from the questionnaire. An explanation of the scales is set out in Table 3.4 in Chapter Three and the results are shown in Figure 5.12 and Table 5.28.

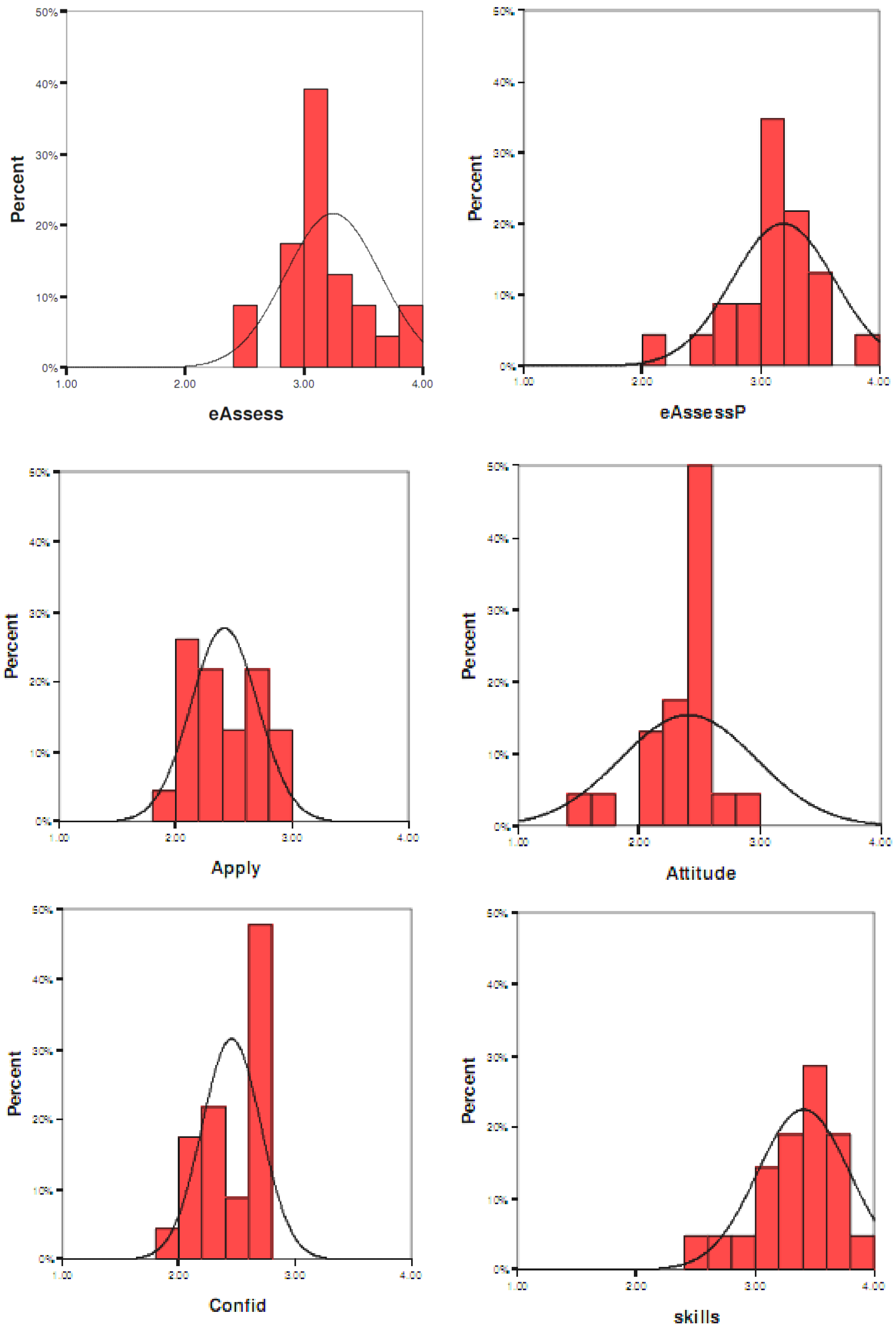


Figure 5.12 XA graphs for the distribution of scores for scales on the student questionnaire

Table 5.28

XA-Descriptions and Descriptive Statistics for Scales Based on Items from Student Questionnaire

	N	Min	Max	Mean	SD	Description
eAssess	25	2.55	4.00	3.24	0.40	Perceived efficacy of computer use for the examination. Potential range between 1 and 4
eAssessP	25	2.18	4.00	3.19	0.43	Perceived efficacy of computer use for the portfolio. Potential range between 1 and 4.
Apply	23	1.83	2.83	2.41	0.28	Application of computer to various uses. Potential range between 1 and 3.
Attitude	25	1.40	3.00	2.40	0.56	Attitude towards using computers. Potential range between 1 and 3.
Confidence	23	1.83	2.67	2.45	0.25	Confidence in using computers. Potential range between 1 and 3.
Skills	23	2.55	4.00	3.40	0.38	Self-assessment of ICT skills. Potential range between 1 and 4.
SCUse	25	0.0	300.0	128.04	81.55	Estimate of time in minutes per day spent using computers at school.

Students supported the view that both the examination and portfolio were appropriate assessment instruments, preferable to pen and paper, allowing them to demonstrate their ability and relatively quick and easy to complete. Means for both *eAssessP* and *eAssess* were 3.2 being approximately 1.5 standard deviations above the scale mid-points. *Attitude* and C*onfidence* scales were also positively skewed with means of 2.4 and 2.5 respectively on a scale of 1 to 3, being approximately one and two standard deviations above the midpoints. Students had a high (self-assessed) level of ICT skills across a range of applications (mean=3.4, midpoint=2.5, SD=0.4). These figures were in line with the population as a whole. On average they indicated using ICT for a little over 2 hours per day at school.

Student Forum

Five students (3 male 2 female) agreed to stay behind immediately after the examination to be part of the student focus group in a semi-structured discussion that ran for 7m 14s.

What did you think of the task(s) you were asked to do?

The students' initial reaction was that the tasks were "pretty easy", "simple to understand" and "easy to follow".This prompted the supplementary question, "Was it demanding enough for a stage 2 exam?" Student all concurred that it was because it “summed up what we have been doing this year”, though all agreed lack of time was a problem with the practical section. Comments were made about the mark allocation and the apparent disparity between marks for design and marks for production. Two students thought that ten marks had been available for planning and were surprised to see it was only five against three for the brochure. The feeling was that the practical should have been relatively more valuable.

Were you able to do your best quality of work?
A chorus of “No!’ greeted this question, with subsequent comments, “you can always do better if you’ve got more time” and “would have been better if the computers were like no complications”

Did the computers help?
Yes. “Definitely easier” but again subsequent comments centred on technical glitches such as pop ups that wouldn’t close or software that hung forcing a re-boot.
How much different was this to how it used to be done?
“Obviously it’s a lot more practical” “It’s more relaxed” (being in a classroom rather than lined up in an examination hall). All concurred that their preference would be for this type of assessment over the written paper, citing “less pressure” and “easier than writing by hand”. When prompted if this was a fairer test most agreed that it was though some students commented that they were better at practical whilst acknowledging others are stronger at theory. “This is how we work in class time so it’s fairer”

What, if anything, would you like changed in future?
“Make it more varied” The student explained that everybody would do the same given the same materials. The reflective questions were perceived as not requiring such a long time. “I don’t see why we had a whole hour to do that. I finished in ten minutes and would have rather had that time to do the logos and brochures and stuff”. “And some of the questions were really awkwardly worded” “like what was your conductivity (sic) and does that mean…it took me five minutes to figure out” (general concurrence from the whole group).

Were there any technical problems with doing the activities?
“Sometimes those computers lag a lot and can freeze” particularly when multiple

applications are open.

In summary, students perceived the examination to be easy to understand and follow, and were positive about having their practical skills assessed. They found the examination to be less pressurised and fairer because it was more in line with what they had been doing in class. They would have preferred more variety in the practical task, thought the time allocated to the reflective questions was excessive and the questions themselves confusingly worded.

Email 'Interview' with Teacher

What did you think of the assessment tasks overall? “Much better as a practical component”

What did you think of the structure of the activities? “Would have liked more preparation time. Delaying the exam does not allow much time for revision etc.”

What were the students' reactions to the activities? “Appeared to be very positive”

What do you think of its potential?
“Excellent. My only concern is with hardware and networks”

What did you think of the quality of work produced by your students for these tasks?
“Did not get to view these which was disappointing because it would have given me an indicator of where skills need to be developed”

Were you surprised by the performance/attitude of any students?
“Positive all round”

What was the general feedback from students?
“They thought the practical exam was harder to complete than they thought the restrictions as to what to do helped keep them thinking”.

Were there any technical problems with implementing the activities?
“Only slow computer – network slow as usual”.

Were there any other problems with implementing the activities? [No response]”

Any other thoughts or suggestions for developing the use of digital forms of assessment? “Earlier release of exam questions”.

In summary, the teacher was very positively disposed towards assessment in the form of a practical examination considering it to have excellent potential with only the worry of systems failure as a downside. Some opportunity to practice this type of assessment under examination conditions would have been welcomed.

Results of Marking

The two external assessors marked five pieces of work for each student comprising the design document for the product, the product itself, two further digital artefacts, a theory section and a practical section of a three-hour examination. The teachers provided scores for the examination but not the portfolio and no semester mark was received. The results for each student are shown in Table 5.29.

The mean of the average of the external assessors’ marks (39.8% with SD 13.5) was very close to the mean for all 115 students involved in the study (36.9% with SD of 17.2%). There was one student in this class (XA118) ranked in the top 10% of students (*Rank Assess Ave*). The mean ranking on the external assessors was 51.1 which was below the mid-point of 58. The range of ranking of Assessor Average- (from 6th to 95th) indicates that this class was diverse in ability and performance. On average, students performed much better on the examination than on the portfolio (examination mean 55%, portfolio mean 35%) but this may have been due to the nature of the portfolio product, an animation. Students were observed to be working with very large files which, when it came to on-line marking, failed to download and display correctly in a web browser and thus could not be properly assessed.

A correlation analysis was done on the rankings and on the scores generated. The results are shown in Table 5.30 and 5.31. Correlation between the external markers total of examination and portfolio scores was strong and significant for both the marking and ranking of students ($r=0.91$, $p<0.01$ and $r=0.92$, $p<0.01$ respectively). There was a

Table 5.29

Results for Case XA from Marking Portfolio and Exam (N=29)

St ID	Ass marking (Total %)			Assessors		Teacher	Rank of	
	Ass 1	Ass 2	Ave	Pfolio (70)	Exam (30)	Exam	Ass Ave*	Tch Exam
xa101	44	50	47.0	34.0	13.0	38	32.0	27.0
xa102	27	29	28.0	15.5	12.5	36	77.5	28.0
xa103	45	45	45.0	28.0	17.0	68	37.5	6.0
xa104	32	34	33.0	17.5	15.5	46	64.5	21.5
xa105	21	19	20.0	12.0	8.0	42	90.0	25.5
xa106	55	62	58.5	37.5	21.0	57	16.5	12.0
xa107	30	22	26.0	9.0	17.0	51	81.0	13.5
xa108	41	31	36.0	18.0	18.0	72	58.0	4.0
xa109	18	21	19.5	5.0	14.5	43	91.5	24.0
xa110	21	30	25.5	9.5	16.0	47	83.5	18.0
xa111	53	60	56.5	37.0	19.5	47	19.0	18.0
xa112	17	18	17.5	a	17.5	60	95.0	11.0
xa113	50	56	53.0	35.0	18.0	74	24.0	3.0
xa114	61	58	59.5	42.0	17.5	69	14.5	5.0
xa115	47	40	43.5	27.0	16.5	49	39.5	15.5
xa116	50	43	46.5	26.0	20.5	85	34.5	1.0
xa117	45	34	39.5	22.5	17.0	62	50.0	10.0
xa118	59	73	66.0	44.0	22.0	64	6.5	8.5
xa119	41	50	45.5	26.5	19.0	46	36.0	21.5
xa120	26	24	25.0	13.5	11.5	47	85.0	18.0
xa121	32	33	32.5	18.5	14.0	51	66.5	13.5
xa122	58	65	61.5	47.0	14.5	46	11.0	21.5
xa123	41	45	43.0	19.5	23.5	64	42.0	8.5
xa124	41	41	41.0	26.0	15.0	49	46.0	15.5
xa125	29	30	29.5	19.0	10.5	42	74.0	25.5
xa126	40	46	43.0	25.0	18.0	75	42.0	2.0
xa127	46	53	49.5	30.5	19.0	46	27.0	21.5
xa128	35	40	37.5	18.5	19.0	67	55.5	7.0
xa129	23	29	26.0	16.0	10.0	31	81.0	29.0
Mean	38.9	40.7	39.8	24.3	16.4	54.3	51.1	
SD	12.8	14.8	13.5	11.0	3.6	13.4	27.1	

*Ranking of external assessors is for all 115 students involved in the study. a Portfolio not submitted.

strong and significant correlation between the teacher's examination score and the assessors' average examination score (r=0.66, p<0.01).

Table 5.30

XA Correlation Coefficients from the Marking of the Student Work (N=29)

	Assessor 1	Assessor 2	Assessor Average	Teacher
Assessor 1 (Total)	1.00	0.91**	0.97**	a
Assessor 2 (Total)		1.00	0.98**	a
Ass Average (Exam)			1.00	0.66**
Teacher Examination				1.00

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

a: no semester marks received.

Table 5.31

XA Correlation Coefficients from the Ranking of Student Work (N=29).

Rank of	Assessor 1	Assessor 2	Assessor Average	Teacher
Assessor 1	1.00	0.92**	0.98**	a
Assessor 2		1.00	0.98**	a
Ass Average (Exam)			1.00	0.73**
Teacher Examination				1.00

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

a: no semester marks received

Conclusions

Conclusions pertaining to this case study are discussed separately for the portfolio and the examination.

Portfolio

Students reported varying degrees of previous experience with portfolios but were strongly positive towards assessment by this method claiming it was quick, easy, good and preferable to alternative methods, allowing their skills to be showcased. The teacher was also familiar with and positively disposed to this type of assessment task which matched the open-ended nature and authenticity of the typical classroom practice. There was a wide variation in the quantity and standard of work submitted for the portfolio. The product, a safety animation, was attempted by a variety of methods and application software. Very few working products were received. Some file sizes were unmanageably large and this impeded remote marking as files had to be downloaded before they could be viewed.

Examination

The computer-supported production examination was implemented with no significant technical difficulties with all students able to complete the requirements of the examination in the time permitted. The students responded well to the style of examination and appeared positively disposed towards the practical component saying that this was more equitable. “This is how we work in class time so it’s fairer”. Many students didn’t seem to appreciate the difference between a logo and a banner or poster. Frequently a logo was made from a photograph by adding text. Numerical data was handled well. Although agreeing that the theory section was made easier by being done on a computer, many students found the reflective questions difficult to understand and repetitive in nature. Many appeared to misunderstand the subtleties of the questions and to which part of the technology process they referred. “And some of the questions were really awkwardly worded” “like what was your conductivity (sic) and does that mean…it took me five minutes to figure out.” Many answers were off topic.

Case Study ZA: Private School

The ZA case study involved one teacher and a class of 17 male, Year 11 students who developed an e-commerce website as the product for the portfolio. The context for the class was Business Information Technology.

Implementation, Technologies and Issues Arising

The researcher either met or communicated using phone and email with the teacher before the students became involved. This was to discuss the process with the teacher and to test the technologies, in particular the use of computers in the school’s laboratory with sound recording and a USB flash drive. The teacher followed the design brief supplied, but this was implemented as an *addition* to other coursework and not as the coursework itself. The performance examination was also undertaken as an additional task, and a difference between this and the other cases was that the two parts were undertaken on different days with only the practical component being observed. The theory section was completed first in a single lesson; the practical component was subsequently completed in two merged periods. Neither the portfolio nor the examination counted towards the student’s final semester mark. The class was conducted in a computer laboratory pictured in Figure 5.13.

Data Collected

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 assessment tasks.

Observations of the Classes

The class was visited on four occasions: to observe students completing the portfolio tasks, to prepare for the examination, to observe and invigilate the examination, to administer the student questionnaire and to conduct interviews with a small group of students about the examination. Computers were up to date and application software extensive including the latest *Office* and *Adobe Creative Suite*.

Figure 5.13 ZA computer laboratory.

Visit 1 Product Development (09/09/08, 1.25-2.20)

Fifteen students (all boys) were present in the computer laboratory. The researcher addressed the class to explain the project structure and participation consent requirements.

They had started the production task (the e-commerce website based on the design brief supplied) having already done the investigation and design phases. Students worked independently for the whole session on the production of graphics or web-pages based on hand drawn screen designs and concept maps. They used *Dreamweaver* or *Fireworks* from the *Adobe Suite*. Many students were involved with developing graphics (e.g.

logos) some used clipart others downloaded images. All those observed appeared to be quite skilful at image editing. Some students were working on the index web-page with one student using a template from *Dreamweaver*. All others appeared to be creating their own theme and layout. The student using the template had the most developed product and indicated that he decided to use the template so he could get more done. A few were observed working on background images and a lot of layering was in evidence.

The teacher gave feedback and suggestions and helped two or three with particular skills. Some students seemed to get ideas from others although on the whole each student's work appeared reasonably original. This was a very productive session with all students on-task for the whole time, particularly on banners, navigation bars, backgrounds, logos and images; very few added much text. Two students appeared to be distracted by *Google Maps* but in fact were on task as they wanted to embed this in their website to indicate how to get to their store.

Visit 2 Development of Design Process Document (20/10/08, 11:00-12:00)

Seventeen boys were present. The class was very attentive, polite and working on the *Miss Shoppe Design Process Document*. Designs had been done on paper and these were being scanned for inclusion. Quite a lot of background research seemed to have been done and some good products (websites) were in evidence. The *Adobe Suite* (*Dreamweaver* and *Fireworks*) had been used skilfully and good documentation was also evident. The teacher was concerned with the amount of time being taken and about file sizes for the portfolio. She suggested using the USB memory at examination time for collection of the portfolio files. The examination was to be an additional task for the students. There was also some concern about access to a *PDF writer* for the examination, it being a requirement to submit files in this format. The teacher handed out a copy of the practice examination to the students who had finished the design document. Students set to work on this immediately and came up to ask for clarification. Other students continued with the *Design Process Document* which was essentially an *after the event* collation of development ideas and post-production reflections.

The examination would be run over class time and thus have to be divided into two sessions. The students would do a *Tertiary Entrance* type of written paper during the examination period. The teacher said that the reflective questions would be scheduled in a single one hour lesson and subsequently, on another day, arrangement would be made to overrun into an adjacent lesson to make time to complete the 2 hour practical component. A headset was left behind in order to check audio recording capability of the computers using the *Audacity* software.

Visit 3 Examination (27/11/08, 8:00- 11:30)

Sixteen students were present in the computer laboratory. The class was just scheduled to do the practical component as they had already completed the reflective questions in a previous lesson. Sound testing was completed successfully using *Windows Sound Recorder*. Initially the teacher organised the class into middle of the lab for the 15 minutes of design work to take advantage of the plentiful desk space. One student opted to plan and design in *PowerPoint* and remained at his workstations. *Adobe Fireworks* was chosen by all for the logo. The brochure was also done with *Adobe Fireworks*. Some good use was made of the supplied images, suitably manipulated, in logos and as defining shapes. Sound recording occurred with what appeared to be much amusement by the students but all functioned without any problems. Students were fully occupied for the two hours

Visit 4 Student Survey and Forum (20/11/08, 7:00-12:00)

All 16 students completed the survey questionnaire. A group of five students was selected by the teacher to take part in a student forum. They were presented with the same set of questions as other cases but follow up questions differed depending on responses. The teacher also remained behind for the student forum and this might have been a constraint on the students who often glanced over towards her.

Survey of Students

Immediately following the examination, 16 students completed the survey questionnaire. The minimum, maximum, mean and population mean (mean for all 115 students across the seven cases) were calculated for each closed response item using SPSS (refer to Appendix K).

Items Concerning the Portfolio

Students reported some previous experience in completing computer-based portfolios with a mean of 2.0 on a scale of 1(Lots of experience) to 4 (No experience). Eleven of the 16 had either lots or some experience and only one responded that he had no previous experience. Students response to the question, *how much more time would you need to get used to it?* (item P1(b)) were split with nine students responding *little* or *no* time, six responding *some* and *one* responding *lots of time* producing a mean of 2.7 on a scale of 1(Lots of time) to 4 (No time). Students also strongly agreed with the statements in items P2(a)-P2(i) which asserted that the computer was *quick*, *easy* and *good* for *developing* and *presenting* their ideas and *creating*, *reflecting* and *demonstrating* skills. Means for these items were all lower than the population means ranging from 1.3 to 1.9 on a scale of 1 (strongly agree) to 4 (strongly disagree). Students' preference for computer-based portfolios was also very strong with responses to items such as *it was better doing the portfolio on a computer than on paper* and *overall the computer is good for portfolios* producing means of 1.3 and 1.4 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to population means of 1.6 for the same items. Only in response to item P2(g), *it was easy to follow the steps to create the portfolio,* was any disagreement apparent with two students disagreeing with the statement; the mean for this item was 1.9 compared to the population mean of 2.0.

Items Concerning the Examination

Students reported little previous experience in completing computer-based examinations with 13 out of 16 responses being 3 or 4 on a scale of 1 (Lots of experience) to 4 (No experience) producing a mean of 3.1 compared to a population mean of 3.2. Responses to the item E1(b), *how much more time would you need to get used to it?,* produced a mean of 2.6 on a scale of 1 (Lots of time) to 4 (No time) with most students responding that some or a little time would be required. No student responded that they would require *lots* of time. Students were very positive about using the computer for the examination, agreeing or strongly agreeing with statements E2(a)-E2(i) suggesting, it made things *quick* and *easy* and was *good* for creating the logo, graphs and brochure and *reflecting* on ideas. Means for these items ranged from 1.3 to 2.1 on a scale of 1 (strongly agree) to 4 (strongly disagree) largely in line with population means. The item, *overall I was able to show what I could do in the exam,* produced a mean of 1.9 in line with the population mean of 1.9 with only four students disagreeing, none strongly. Overall, students' preference for a computer-based examination was strongly supported with the statement, *it was better doing the exam on computer than on paper,* producing a mean of 1.6 on a scale of 1 (strongly agree) to 4 (strongly disagree) compared to 1.7 for the population as a whole. Only two out of 16 students disagreed with this statement

Items Concerning Computer Use

Students indicated that they were regular and extensive users of ICT with all 16 reporting computer use at home on a daily basis. All but two students had home broadband internet and a full range of home entertainment and communication systems were reported; 14 owned a mobile ‘phone. Attitude towards, and confidence with computers was definitely positive with 15 out of 16 students responding *yes* to the statements, *I'm good at using computers* and *I feel confident working on computers*. Only one student felt that computers were not *good for the world*. Students' self-assessment of their proficiency with application software was varied. Whilst high ability was indicated with word processing, slideshows, email, file management, internet research, digital photography and image editing (means 3.4 to 3.9) students were less confident with spreadsheets (mean=3.2), databases (mean=2.5) and digital video (mean=2.6) on a scale of 1 (low ability) to 4 (high ability).

Open-ended Items

Responses to the open-response items were tabulated to assist in drawing out themes and a sample of these are presented in Table 5.32. Generally students considered that using computers made it easier, better, quicker and more fun, and also provided a better environment within which they could use their skills and demonstrate their ideas. The main worst things were time pressure in the examination and to some extent the portfolio and the usual concern that the computer could crash and their work might be lost. Interestingly for this group, the possibility of copying or being copied from was mentioned as one of the two worst things about the examination the perception being that this was a disadvantage rather than an advantage. Several responses to the two worst things about both the portfolio and the examination were left blank.

Questionnaire Scales

Seven scales were derived from combining selected items from the questionnaire. An explanation of the scales is set out in Table 3.4 in Chapter Three and the results are shown in Figure 5.14 and Table 5.33. Students strongly supported the view that both the portfolio and the examination were suitable and appropriate assessment instruments, preferable to pen and paper, allowing them to demonstrate their ability and relatively

Table 5.32

ZA Summary of Students' Responses to Four Open Ended Questions on Survey Questionnaire

The two best things …	The two worst things …
Portfolio	Portfolio
The predominant theme *easiness* was mentioned in 9 out of 32 responses in slightly different contexts as exemplified by "Easy to follow" and "Easy to show skills" and "I found it was easier to have it on the computer than the hardcopy" and "It was easier to convey ideas" and "It was easy to fix errors". There were three references to *fun* e.g. "Was fun" and several to various degrees of superiority for example "I could show my skills on the computer" and "Better for design" and "Structured and organised".	A main theme was the danger of lost or corrupted files, for example "Computer error messed up data" and "If the folio becomes corrupt you can lose all your info". Four responses mentioned *time* for example "Time consuming" and "Not enough time". Six out of 32 responses were blank.
Examination	Examination
Easy, Quick and Comfortable (8 students made reference to easy or easier, 3 to quicker) e.g. "Easier to show artistic capabilities" and "Quicker to type" and "It's a familiar place"	Time pressure (9 students lamented the lack of time to complete the practical task) e.g. "Limited time for complex techniques" and "Not enough planning time"
Enabling (5 students made reference to the suitability of the task) e.g. "Able to show Practical skills that we have been practicing over the year" and "I was able to demonstrate skills I have learnt".	Hardware failure (5 students mentioned the potential for disaster). e.g. "Worrying the computer will stuff up"
Enjoyable (3 students referred to the exam as fun or enjoyable) e.g. "Wasn't as boring, was doing something that's fun" and "Enjoyable task".	Proximity (3 students alluded to potential implications of sitting close together) e.g. "It is too easy to copy someone else's design" and "You're right next to your peers"

quick and easy to complete. The mean for *eAssessP* was 3.4, more than three standard deviations above the midpoint; the mean for *eAssess* was 3.3 almost three standard deviations above the scale mid-points.

Table 5.33

ZA-Descriptions and Descriptive Statistics for Scales Based on Items from Student Questionnaire

	N	Min	Max	Mean	SD	Description
eAssess	16	2.82	4.00	3.32	0.34	Perceived efficacy of computer use for the examination. Potential range between 1 & 4
eAssessP	16	3.09	4.00	3.40	0.25	Perceived efficacy of computer use for the portfolio. Potential range between 1 & 4.
Apply	16	2.00	3.00	2.57	0.30	Application of computer to various uses. Potential range between 1 and 3.
Attitude	16	2.40	3.00	2.73	0.24	Attitude towards using computers. Potential range between 1 and 3.
Confidence	16	1.00	3.00	2.72	0.53	Confidence in using computers Potential range between 1 and 3.
Skills	16	2.27	3.91	3.39	0.46	Self-assessment of ICT skills. Potential range between 1 and 4.
SCUse	16	50.0	192.0	82.93	39.4	Estimate of time in minutes per day spent using computers at school

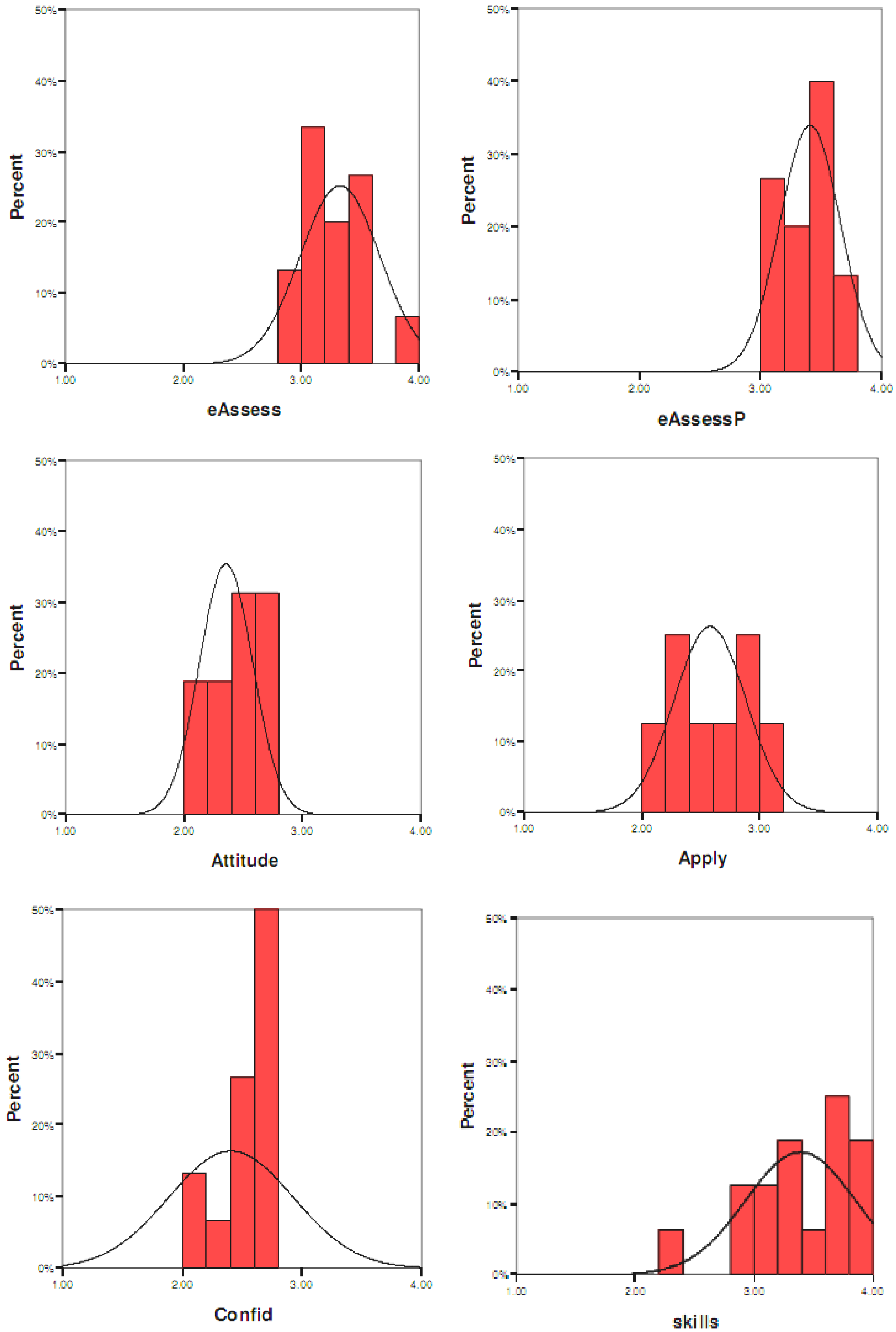


Figure 5.14 ZA graphs for the distribution of scores for scales on the student questionnaire

Attitude and *confidence* scales were also positively skewed with means of 2.7 on a scale of 1 to 3, being approximately three and two and a half standard deviations above the midpoints. With the exception of databases, students had a high (self-assessed) level of ICT skills across a range of applications, such as spreadsheets, digital photography and web authoring (mean=3.4, midpoint=2.5, SD=0.5). On average used ICT (*SCUse*) for more than one hour per day at school.

Student Forum

Four male students (this was an all boys' school) agreed to stay behind immediately after the examination to be part of the student panel. The class teacher chose to be present and this did appear to constrain responses. The discussion went on for 5m 26s.

What did you think of the task(s) you were asked to do?
"The tasks were well explained and easy to complete".

"Similar to a task we've already done". "A good change from just writing it up...so you had more opportunities to show your skills". "Pretty good because you're put under time pressure like in a real life situation". "A bit hard to understand the words [in the reflective questions]".

Did the computers help?
"With the time restraints I don't think you could [produce your best work]" "With more variety of software we'd have been able to do better stuff"

How much different was this to how it used to be done?
"It's totally different and it's a lot better because we spend the whole year on computers and then suddenly we're put in a room with a pen and paper...it doesn't really suit" "I think this idea is a lot better" "I found with the designing part [the logo] you were able to use more effects" "With the computer you can show effects like transparency and mixed colours."

What, if anything, would you like changed in future?
"More time to do it." "More reliable computers."

Were there any technical problems with doing the activities?
“Some programs…it lagged a bit” “Some programs don’t work with others...like you can’t cut and paste in…I had to use print screen and work my way around it”

In summary, students welcomed the practical format of the assessment and the opportunity to demonstrate their creative skills. All were in favour of a practical assessment agreeing that computers were integral to the course and therefore an essential part of the assessment process. Students would have like more time to produce better work but recognised the authenticity of time pressure.

Email ‘Interview’ with Teacher

What did you think of the assessment tasks overall?

> I felt that the task was excellent-There was a lot of scope in terms of design with the challenge of a logo and then a website-As my students were boys I went with the Miss Shoppe as I felt it would challenge them, and it did. As well as specifying pages for the website a merchandise page should also have been included.

What did you think of the structure of the activities?

> I felt that there was not enough time for development of the website-The students wanted to do a really good job, and felt frustrated that they could not complete all pages to a satisfactory standard. The instructions were too wordy, and often repetitive, very often they did not understand the questions in the reflective section of the document. They needed to scaffold their response in the format of a word report detailing their progress, and responses-I believe most would have done quite well at this.

What were the students' reactions to the activities?

> Positive, although they felt under pressure, and we had real problems with the school network playing havoc with their files, and subsequently we lost work, and access to computers on at least three occasions. The exam was ok, although creating a brochure was not what we really wanted to do; I felt there should have been a choice of an interactive product as well. The students would really have been able to showcase their skills in designing web

interactive elements rather that a published document-This is what they wanted to do.

What do you think of its potential?

Yes it has some validity; however it must work in conjunction with a digital portfolio, as well as a task. You cannot judge a student’s ability alone based solely on work in an individual task. Instructions on how to collect relevant work must be very clear, and less verbose-the kids couldn't get their head around ‘digital artefact 1’-why not call it Student Work Sample One. In other words simple clear language.

What did you think of the quality of work produced by your students for these tasks?

Students produced basic work in exam. It would have been better to have a choice of ICT products. Their performance in the task brief was not indicative of their ability when compared to semester's work, as they had more time, and were able to invest time in researching and troubleshooting problems with scripts, code etc..

Were you surprised by the performance/attitude of any students?

They were pleasantly pleased about the task; however as a teacher trying to get through the syllabus, it was way too onerous. I did not cover syllabus content in year 11 AIT, as I was pushing to get this task completed. This made me very anxious, and grumpy!

What was the general feedback from students? (would they like more of it?)

Unhappy that they did not have enough time-preference for a choice of products-computers too slow-desire for a more simplified set of task instructions. They also wanted the opportunity to have a written exam as well, as some felt that they do very well in communicating concepts than actually designing them.

Were there any technical problems with implementing the activities?
“Just school computers and downtime which affected us severely”.

Were there any other problems with implementing the activities?
"Time-Time-Time! I did not cover syllabus content in year 11 AIT, as I was pushing to get this task completed. This made me very anxious, and grumpy!"

Any other thoughts or suggestions for developing the use of digital forms of assessment?
"The marking key needs a serious revamp-You cannot drill down into fine detail and award marks-it also reads like an outcomes based marking key. I had to write my own, and this was disappointing, but more consistent with other submitted tasks".

In summary, the teacher commented positively on the tasks in both the examination and the portfolio, describing them as excellent particularly in their open ended nature. Time constraints were cited as detracting from students' performance on the portfolio tasks. The examination was perceived as too restrictive and overly complicated. The failure of systems also adversely affected completion of the assessment tasks.

Results of Marking

The two external assessors marked five pieces of work for each student comprising the design document for the product, the product itself, two further digital artefacts, a theory section and a practical section of a three-hour examination. Due to a *server meltdown* the teacher was not able to provide separate scores for the portfolio or the examination but did provide a semester score. The results for each student are shown in Table 5.34 The mean of the average of the external assessors' marks (53.6%) was statistically significantly higher than the mean for all students involved in the study (36.9% with SD of 17.2%). There were six student in this class ranked in the top 10% of students (*Rank Assess Ave*). The mean ranking of the external assessors was 27.5 which was well below the mid-point of 58. This class achieved well above average with most students doing well in spite of the fact that the portfolio and the examination were not part of their semester assessment programme.

A correlation analysis was done on the rankings and on the scores generated. The results are shown in Table 5.35 and 5.36. Correlation between the external markers was strong with a correlation coefficient of 0.78 ($p<0.01$) for the marking, and 0.86 ($p<0.01$) for the ranking of students based on their total scores. There was little or no correlation

Table 5.34

Results for Case ZA from Marking Portfolio and Exam (N=16)

St ID	Ass Marking (Total %)			Assessors		Teacher (%)	Rank	
	Ass1	Ass2	Ave	Pfolio (70)	Exam (30)	Sem	Ass Ave*	Tch Sem
za101	68	62	65.0	40.5	24.5	78	9.0	3.0
za102	70	62	66.0	47.0	19.0	80	6.5	2.0
za103	41	37	39.0	28.5	10.5	71	51.5	7.5
za104	35	51	43.0	29.5	13.5	51	42.0	16.0
za105	53	62	57.5	39.0	18.5	53	18.0	15.0
za106	68	80	74.0	52.5	21.5	75	2.0	4.0
za107	33	20	26.5	17.5	9.0	70	79.0	9.5
za108	45	36	40.5	25.0	15.5	61	47.5	12.0
za109	81	64	72.5	50.5	22.0	84	3.0	1.0
za110	83	66	74.5	51.0	23.5	72	1.0	6.0
za112	54	50	52.0	34.0	18.0	57	25.0	13.0
za113	57	65	61.0	43.0	18.0	74	12.0	5.0
za114	45	52	48.5	32.0	16.5	70	28.5	9.5
za115	32	28	30.0	14.0	16.0	71	72.5	7.5
za116	49	48	48.5	31.0	17.5	62	28.5	11.0
za117	54	65	59.5	37.5	22.0	54	14.5	14.0
Mean	54.2	53.0	53.6	35.8	17.8	67.7	27.5	
SD	16.1	16.0	15.2	11.5	4.4	10.1	24.6	

*Ranking of external assessors is for all 115 students involved in the study.

Table 5.35

ZA Correlation Coefficients from the Marking of Student Work (N=16)

	Assessor 1	Assessor 2	Assessor Average	Teacher Semester %
Assessor 1	1.00	0.78**	0.95**	0.52*
Assessor 2		1.00	0.94**	0.17
Average			1.00	0.36
Teacher Semester %				1.00

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

Table 5.36

ZA Correlation Coefficients from the Ranking of Student Work (N=16)

Rank of	Assessor 1	Assessor 2	Assessor Average	Teacher Semester %
Assessor 1	1.00	0.86**	0.95**	0.39
Assessor 2		1.00	0.97**	0.14
Average			1.00	0.28
Teacher Semester %				1.00

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

between the scores awarded by the assessors and the teacher's semester scores even though both rated some of these students as among the best in the population.

Conclusions

Conclusions pertaining to this case study are discussed separately for the portfolio and the examination.

Portfolio

Students reported some previous experience of portfolios and were divided over the amount of time required to become familiar with this form of assessment. Students were however positive about the portfolio, claiming it was a quick, easy and fair method of assessment allowing them to demonstrate their abilities and preferable to written assessments. The quantity and standard of work submitted for the portfolio were well above the population average. The product, a website, was attempted using the same application software (*Adobe Creative Suite*), and work was generally of a high standard with full use of the software's capability demonstrated.

Examination

The computer-supported production examination was implemented with no significant technical difficulties. Several students reported time pressure. The quality of work was overall very high and students attempted some complex effects which might have impacted on available time. Students responded well to the style of examination and appeared positively disposed towards the practical component saying that this was a more suitable form of assessment. "It's a lot better because we spend the whole year on computers and then suddenly we're put in a room with a pen and paper…it doesn't really suit" These students felt somewhat limited by the constraints of time and the design brief. Numerical data was handled well. Although agreeing that the theory section was made easier by being done on a computer, many students found the reflective questions difficult to understand failing to pick up the subtleties of the questions and to which part of the technology process they referred. "A bit hard to understand the words [in the reflective questions]" Evidence for this were the many off topic answers, though in general this group performed well in both theory and practical activities.

Summary

This chapter has described in detail the implementation of the portfolio and the examination at each of the seven schools involved in the study. The implementation of the portfolio was inconsistent across cases. The differences between what was intended and what actually occurred were sometimes large and sometimes small. Each school differed in some regard to the others, for example, in the nature of the product or how the portfolio counted, if at all, towards the overall semester mark. However, the questionnaire and questionnaire scales revealed a consistent and highly positive attitude towards the portfolio as a suitable, fair and appropriate alternative assessment instrument. Some idea of the variation in implementation of the portfolio task may be ascertained from Table 5.37

Table 5.37

Variations in the Implementation of the Portfolio assessment Task

Aspect	CA	LA	MA	RA	WA	XA	ZA
Design brief followed without modification	No	No	Yes	No	No	No	No
Lesson time allocated as set out in design brief	Yes	No	Yes	Yes	Yes	Yes	No
Assessment included in students final mark	Yes	No	Yes	Yes	No	Yes	No

Of the eight classes in seven schools, only one (MA) implemented the portfolio exactly as intended, using the supplied task design brief without modification, allocating the time specified and including the assessment of the task as part of the students' overall semester mark.

The examination, being highly structured and time bound, was able to be implemented with far greater consistency than the portfolio. Only in the first case (RA) was there a major difference; the sections were completed with the theory section preceding the practical. Although there were a few technical difficulties, these did not prevent the task from being completed fully by all students, with the exception of the audio recording section which had to be omitted in some cases. Responses to the questionnaire and in the student forums indicated a highly positive attitude towards the idea of a computer-based assessment involving a practical component. Students considered the examination to be superior to the current pen and paper system, being easier, more appropriate to the course and in line with their day to day classroom practice, allowing them scope to demonstrate skills which might otherwise remain unexamined.

In principle, the teachers were positive about both the portfolio and the examination, with some caveats, for example the complexity of the wording in the portfolio and the marking key in the examination. The following chapter brings together the findings from all the case studies, developing and eliciting cross case comparisons in order to present a discussion of the results in terms of the research questions.

CHAPTER SIX

DISCUSSION OF RESULTS

Introduction

This chapter presents a discussion of the results of the study in relation to the research question and subsidiary questions. This discussion will lead to a set of definitive conclusions in the final chapter. The study set out to address the following research question:

How may the digital capture of students' performance most effectively support summative assessment in the senior secondary Applied Information Technology course?

The discussion was framed around the subsidiary questions and then brought together towards the end of the chapter around the overarching research question. The subsidiary research questions were:

1. What were the advantages of digital capture of students' performance in support of summative assessment of practical ability in the senior secondary AIT course?
2. What were the limitations of digital capture of students' performance in support of summative assessment of practical ability in the senior secondary AIT course?
3. How feasible was the digital capture of students' performance in different forms of summative assessment in AIT with respect to
   i) Manageability,
   ii) Technical facility
   iii) Functionality, and
   iv) Pedagogy?
4. Did judgements by multiple comparisons of pairs, produce reliable scores when applied to summative assessment of practical performances in the senior secondary AIT course?
5. Would multiple comparisons of pairs be a better method than analytical marking

of student practical performance in AIT?

Each of these subsidiary questions will now be addressed in turn followed by a general discussion of the results in terms of the overarching research question.

Advantages of Digital Capture

The first subsidiary research question centred on the advantages of digital capture of student performance for a form of summative assessment in the AIT course. Two main forms of assessment were investigated: a digital reflective process portfolio and a computer-based examination. This section will discuss three principal advantages.

The first major advantage of the digital capture of students' practical performance in AIT was the opportunity afforded to students to be assessed comprehensively and authentically. In its absence, assessment would be, and has been, limited and confined to those aspects of the course which could be measured in traditional pen and paper examinations. This places undue emphasis on abstract knowledge of the design process, design principles and conventions, documentation and computing theory, excluding any assessment of the student's practical capability or application of theory to complex problems. Whilst these aforementioned aspects of the course are important and may be suitably assessed by traditional methods, the exclusion of all other course outcomes from the assessment process is likely to lead to several undesirable consequences, not least of which is the marginalisation of practical skills. This is in spite of the perception by teachers and students alike that these skills are the primary reason for the existence of the course. Three of the four AIT course outcomes, *Technology process, Understanding information and communications technologies and Quality of information solutions,* clearly specify the value and importance of practical skills (Curriculum Council of Western Australia, 2009).

The tasks developed for the portfolio and the examination were built on the course outcomes and content and were intended to be authentic in both their fidelity to these, and to common or typical real world applications of technology. Table 6.1 shows extracts from the course outline together with some examples of their instantiation within the portfolio and examination. The tasks for the Portfolio and Examination were perceived very favourably by teachers with six of the seven describing them as appropriate or excellent in their responses when surveyed with an e-mail questionnaire.

The following three quotes illustrate the perception of the teachers:

> The assessment tasks were really good. (CA)
>
> I thought the tasks were excellent, easy enough for less able students to achieve an adequate result but open ended enough for the good students to excel. (WA)
>
> I felt that the task was excellent-there was a lot of scope in terms of design. (ZA)

Table 6.1

Alignment of Practical Outcomes between Course Requirements and Digital Assessments.

Course Outcome	Digital Portfolio	Computer-based Examination
…*implement* and evaluate production processes and strategies to manage resources efficiently.	Creation of website in extended time frame. Reflection and appraisal of this digital solution.	Creation of marketing brochure in short time frame. Reflection and appraisal of this digital solution.
…understand the nature and *use* of computer hardware and software to achieve information solutions.	Selection and use of suitable application software from a variety of choices to create features of a webpage such as rollover buttons, hyperlinks, hotspots, animations and effects. Creation of two digital artefacts showing breadth of skills.	Selection and use of suitable application software from a variety of choices, to create graphs, charts, logo and brochure and manipulate images.
…*select and use* appropriate software and hardware to achieve information solutions;	Selection and deployment of web authoring application software to create a high impact site with appeal to a niche market.	Using a spreadsheet to collate, organise and present raw data in graphical form with high visual impact.
…*use* skills, techniques, processes, standards and conventions to achieve information solutions.	Creation of a fully functioning home page balanced and colour coordinated with a professional look and feel.	Creation of a logo to suggest low environmental impact by appropriate application of shapes and colours.
…*apply* appropriate forms, structures and conventions to create or modify information solutions;	Create a home page with intuitive navigation, balance and uncluttered look.	Create a themed brochure with visual appeal and uncluttered layout.
…*apply* skills, techniques and processes to develop information solutions; and	Demonstrate skill acquisition by creation of portfolio product and additional artefacts. Creation of process document.	Demonstrate a variety of acquired skills by creation of product featuring original graphical logo, manipulation of numerical data, text, and graphics.
…*apply* enterprising capabilities, exploring alternatives, in working to achieve information solutions.	Creation of an e-commerce shop front to promote and advertise a business case... Creation of process document.	Creation of a marketing brochure to promote and advertise a business case.

Teachers also reported their impression of student perceptions and these too were generally positive with regard to the nature of the tasks, given that they were for assessment purposes as exemplified by the following three quotes.

> They were happy enough with the practical but I got a feeling they weren't all that keen about a practical exam. (CA)
>
> Positive, although they felt under pressure. (ZA)
>
> The students enjoyed the portfolio task. Students liked the idea of the practical exam. (MA)

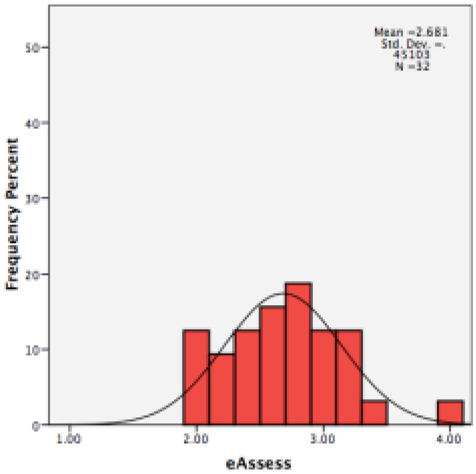
Students also indicated, by survey responses and comments expressed in student forums, that they were positively disposed to the tasks as reported in summary in Chapter Four and in detail, case by case, in Chapter Five. Specifically for the AIT students, 89% either agreed or strongly agreed with the assertion of survey item E2(a) that *It was easy to use the computer for doing the exam* and 86% either agreed or strongly agreed with the assertion of item E2(k) that *overall, it was better doing the exam using a computer than on paper*. These figures were supported by the responses from the student forums where comments such as “the tasks were well explained and easy to complete” and “this is much better…a lot simpler…practicals let you show what you can do…I'm not great at theory” were typical of the positive attitude towards digital forms of assessment. For the portfolio response to items P2(a) that *It was easy to use the computer for doing the portfolio* and P2(b) *It was easy to use the computer for my portfolio to develop ideas* was less enthusiastic with 58% and 49% respectively either agreeing or strongly agreeing. However responses to the other items P2(c)-2(k) were much stronger with a mean of 88% agreeing or strongly agreeing with the assertions. For example 90% of respondents either agreed or strongly agreed with Item P2(k) that *overall, it was better doing the portfolio using a computer than on paper.*

Two scales (*eAssess* and *eAssessP*) developed from a *subset* of questions on the student questionnaire, measured the perceptions of students towards the efficacy of the practical examination and the digital portfolio. Both scales had a mean value of 3.2 on a scale of 1 to 4 where 1 represented strongly negative and 4 represented strongly in favour. These means were well above the mid-point of the scale at 2.5, being 1.6 and 1.5 standard deviations above the mid-point respectively. Responses in AIT may be compared to those of students in three other senior secondary courses (Italian Studies, Engineering

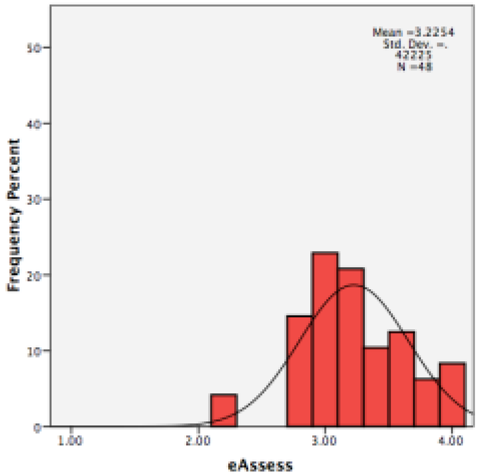
Studies and Physical Education Studies) which also have large practical/performance components and for which concurrent research was undertaken. Figure 6.1 compares the distributions on this scale between all four courses. Note that for the AIT course there are two graphs; one for the examination and one for the portfolio corresponding to the separate questions about each in the questionnaire. Each of the other courses three courses had only one form of assessment. Means for both AIT (portfolio and examination) and Engineering students were very high pointing to a highly positive perception of the assessment method and tasks. For Italian Studies and Physical Education Studies, the scale means were lower indicating a less favourable perception.

The second major advantage of digital capture of performance in AIT was the convenience afforded by digital storage and transmission of student work. The beneficiaries were not only the researchers and assessors; teachers and students were able to back-up and store, collate and transfer their work with ease. Everything that students produced on the computer was already in digital form and those aspects of the assessment tasks such as design sketches, more suitably developed on paper, were easily digitised. Indeed, the creation of digital files and their conversion between multiple formats are essential skills in AIT. No restrictions were placed on the applications used; however the nature of the tasks led most students to use very similar combinations of basic office and multimedia software. Conversion of files to an application independent format such as portable document (.pdf) was a requirement of the examination and most participants had the capability, if not the knowledge, to achieve this. Once digitised, students work was easily and rapidly transferred between locations using portable storage devices or network hardware and secure, reliable, well established protocols.

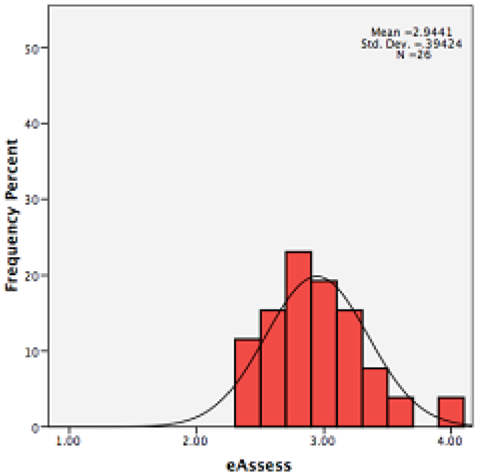
The third major advantage of digital capture of student work follows from the ubiquity of internet access and the expansion and improvement of broadband Internet services. Once the work had been uploaded to the web-based file server, it became a simple matter for assessors to view files, even when these were relatively large, provided a robust internet connection was available. Online marking tools simplified the scoring process by presenting marking keys and rubrics together with the work sample; totalling by computer of scores awarded led to time saving and scoring accuracy. Access to students' work was unconstrained by time or location and assessors did not have to



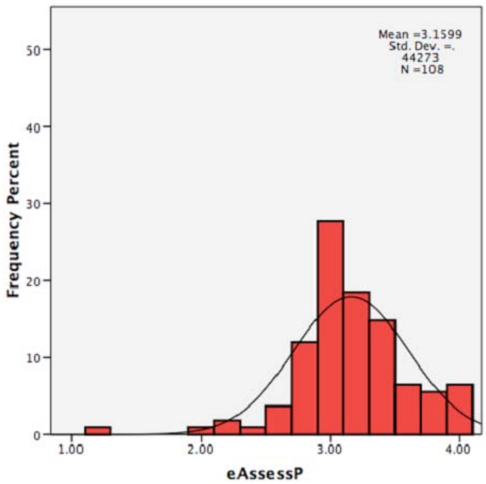
Italian (N=32 Mean 2.7)



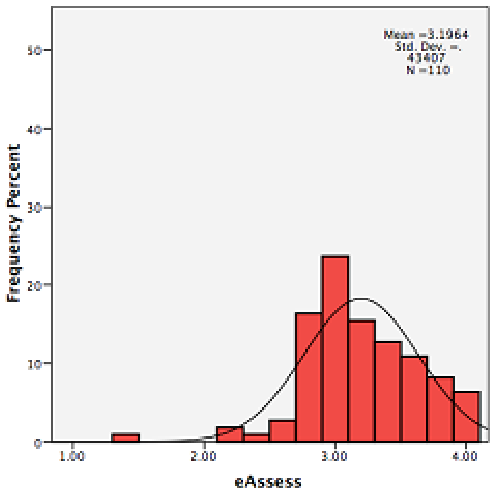
Engineering (N=48 Mean 3.2)



Physical Education Studies (N=26 Mean 2.9)



AIT Portfolio (N= 108 Mean 3.2)



AIT Examination (N= 110 Mean 3.2)

Figure 6.1 Student perceptions of the efficacy of digital assessment in AIT, Italian, Engineering and PE Studies.

travel to a central marking location. Notwithstanding issues of database concurrency, assessors were able to score work and save results simultaneously as illustrated by the following response.

> I was able to complete some of the analytical marking whilst on holiday in Malaysia using one of many open wireless access points. Download speeds were a problem particularly when large files were being viewed and this could have led to error- an unobtainable file might have been assessed as missing and hence awarded a zero score. (Assessor 1)

In summary, three major advantages of digital capture of student performance for assessment were apparent in this study. First and foremost was the opportunity to present assessment tasks which were more closely aligned to the AIT course, allowing assessment to be authentic in nature and more comprehensive in scope with improved face and content validity. Secondly, digital capture allowed synergies in cost, transport, storage, back-up and conversion of digital files leading to improved efficiency of the assessment process. Finally, web-enabled marking tools allowed students' work to be accessed and scored at any time and from any location provided that an adequate internet connection was available.

Limitations of Digital Capture

The second subsidiary research question focussed on the limitations of digital capture. In this section the principal limitations of digital capture of student performance in summative assessment for the AIT course are examined.

The main disadvantage of digital capture brings to the fore an analysis of what it intends to replace; paper based examinations. The advantages of the latter are self-evident; they are tried and tested, easy and cheap to organise, can be done under identical conditions across multiple locations and, most importantly, have widespread public confidence. Computer-based assessment on the other hand, is something new, requires specialised equipment and software, must be conducted in a specialised location, is difficult to standardise and subject to unlikely, but nevertheless possible, constraints such as mains power and equipment failure. For many school administrators, the easy option is to stick with paper based assessment.

The fear of something going wrong and work being lost was commonly expressed by students in the survey and student forums, as reported in Chapter Four, and it remains the major factor working against the adoption of digital capture in courses such as AIT. The impact of such “negative critical incidents” (Kinshuck et al., 2008) were a major factor affecting satisfaction as illustrated by the following students’ comments with respect to the examination.

> That’s my only downside using the computers … if something screws up you don't have anything else. (CA)
>
> I had to shut down programs three or four times… my page closed before I'd saved. (CA)
>
> Most of the programs at the start wouldn't even download. (WA)
>
> Needs to be all set up beforehand to make sure. (WA)

Because the portfolio ran over an extended time period, it did not engender the same level of anxiety as the examination as work could be backed up or deferred to another day. During the examination however, there was widespread failure during the audio recording section but apart from this, problems were few and immediately fixable. For example at CA, three students had to restart their computers when they stopped responding. However for all the other cases reliability of technology did not appear to be a limitation. The details of these problems are discussed under the Technical Facility section of the Feasibility Framework later in this chapter.

A further limitation was connected with the concept of equity. With hardware and software being provided by the school, there exists a real possibility that students with access to more up-to-date hardware could have an advantage. For example, many multimedia applications require extensive system resources to function effectively. Machines lacking in main memory or with slower processors often struggle to run these applications efficiently and may be subject to delays and hangs. This was a factor at one of the state schools (RA) where although the range of software available was extensive, it was delivered from a central server and this did cause minor disruption during the examination. This did not appear to be an issue at the other six schools.

The range of software available to students is likely to vary considerably and this might impact upon the quality and sophistication of student work although all students in the study were able to respond to all tasks. In fact, the relevant software available at all seven schools was very similar being based on the Windows XP operating system, *MS Office* (either 2003 or 2007) and *Adobe Creative Suite*. Students were asked to convert some files into common, application independent formats for assessment and this was possible at all schools and simplified the marking process.

During the examination, none of the participants employed web based application software and resources though these could easily have been used where internet access was unrestricted. Only at two schools (CA and MA) was access to the internet blocked during the examination. At the others, it would have been possible to upload, edit and download one of the supplied photographs using for example *Adobe Photoshop Express* and to create and manipulate a spreadsheet and word process a document with *Google Docs*. The question of access to the internet might therefore be considered a limitation though it would have been unusual for students to use web applications when local equivalents were available. Of course the Internet was available to students for research during the portfolio; in the examination, where available, students were not observed to make any use of it.

Though the examination and portfolio both required specialised locations and equipment in the form of computer laboratories, observed differences between locations were minimal. All schools ran similar operating systems and software suites. Hardware varied in age and specification but the assessment tasks placed no insurmountable burdens on even the least up to date systems, except where students attempted to run too many applications simultaneously.

Feasibility of Implementation

The third subsidiary question focussed on the feasibility of implementation. A summary of findings was compiled from the seven AIT case studies, based on the Feasibility Framework developed from Kimbell et al. (2007) and comprising the dimensions shown in Figure 6.2.

Manageability	Technical	Functional		Pedagogic	Constraints	Benefits
		Validity	Reliability			

Figure 6.2 The Feasibility Framework after Kimbell et al. (2007)

The findings were allocated to the dimensions of Manageability, Technical facility, Functionality and Pedagogy. Each aspect included a summary of the constraints and benefits of the form of assessment used in the context of the specific case. These findings are now discussed in turn with respect to the terms of each dimension of the feasibility framework

Manageability Dimension

Manageability in the context of this study refers to the practicalities of administration, collection and assessment of student work in digital files for both the portfolio and the examination. Likely factors impacting on manageability might be sufficiency or insufficiency of working time and space, opportunity to compromise assessment regulations or the requirement that all work should be original and the student's own. The ease of collection of student work and consistency of teachers' interpretation and administration of the tasks were other potential factors which were examined under this dimension. A summary of some of the main findings for each case study are presented in Table 6.2.

Table 6.2

Manageability Findings from Feasibility Framework by Case

Case	Manageability Portfolio	Manageability Examination
CA	Portfolio managed by teacher with work output submitted on one DVD for class. Engagement of students appeared variable. Teacher adapted task and students understood requirements. 11 of 20 submitted incomplete portfolios defined as at least one section missing	All students submitted work on 4GB USB flash memory. Internet access disabled. Exam manageable though class sizes made for cramped and close seating in exam. Students could easily see neighbouring screens so security could have been compromised. Teacher was happy with the resources provided (practice exam and exam itself).
LA	Portfolio easily managed by teacher with small class size in purpose built lab. Although well understood and explained by teacher, not included in semester assessment hence engagement/ time allocated reduced. Additional artefacts not submitted by 8 of 10 students. Submission of large video files impractical for web based assessment.	Easily managed with small class size in purpose built lab. Students well-spaced so couldn't see adjacent screens however tiered arrangement meant those behind could see those in front. Completed as additional task and not included in semester mark. Nevertheless students tackled exam with enthusiasm and produced good work.
MA	Small class. Teacher and class complied exactly with task requirements. Well explained and understood by students. All students able to complete portfolio -only 2 artefacts missing. Hardware, software, time all sufficient.	Small group easily managed and well-spaced in computer laboratory. Ran smoothly Good hardware and software. 5 mins too short for audio section otherwise, software, time all sufficient. 3 USB s corrupted but re-imaged in less than a minute.

(continued)

		No disruption but students were a little put out-created additional exam stress.
RA	Year 12 students hence reduced time to complete portfolio. Students understood requirements however many portfolios incomplete particularly digital artefacts (9/14 missing. Large variety of software available from central server but delivery observed to be and reported by students to be unpredictable.	First implementation of examination and problem with timing/invigilation hence decision to swap sections to do practical section first. Room and spacing adequate. Two or three crashes (server delivered applications stopped responding) but spare computers available and minimal disruption though extra stress for students. Students repeatedly reminded to save work.
WA	Teacher managed small group in spacious facilities. Portfolio task chosen by teacher involved advertising for real business. Students understood requirements and addressed problem with wide range of solutions, some of which very sophisticated/ over ambitious unable to complete in time available. No student portfolios were received. Some student collaboration observed during visits in problem solving animation with Adobe Flash.	Easy to manage small groups in spacious facilities. Surprisingly, some students reported missing software and teacher commented on network management issues; all had been OK previous day. Software access problem meant examination had to be split across two rooms by opening movable partition. Moved group allowed additional 10 minutes to complete. Teacher concern for cheating .
XA	No difficulties in delivery of portfolio. Teacher chose alternative task. This led to some file management problems with large files being handled and saved to network storage.	Class list inaccurate led to last minute appearance of three additional students. Large group spread across two labs. Could see neighbouring screens and hence student collaboration could not be discounted.
ZA	Teacher managed portfolio and attempted implementation in parallel with her own course. Activities done as extras and practice hence claims of time pressure though all three components submitted for all students. Teacher tried to comply fully. Problem with submitting student work after network problem. Files eventually recovered and submitted.	Split: theory and practical done on separate days to fit in with class timetable. Teacher and researcher invigilated. Some confusion over requirements. Done as extra task and not assessable. Students nevertheless appeared to be fully engaged and delivered some of the better quality work observed. Room allowed good separation between students so student collaboration unlikely.

Considerable flexibility was allowed in the setting up and implementation of the portfolio and it was this freedom which inadvertently introduced the main manageability issues. Schools were selected on the basis of teachers' response to an expression of interest and only when selected, fully involved and committed did the scope of the requirements become apparent to some. Only one teacher (MA) fully complied with all the requirements. Here, students' completion of the portfolio was part of the assessment for the semester and hence sufficient time was allocated to explain the requirements and allow for all parts to be concluded. In other schools, requirements were not fully appreciated by the teacher until the researcher visited. It was a delicate task to attempt to realign the portfolio to fit it into an already full programme.

It is fair to say that the instructions for the portfolio task were clear but that these were either not studied closely or disregarded by some teachers. The freedom extended in the choice of portfolio product, where schools used the e-commerce site supplied, a safety animation, a cinema advert and their own choice of website, appeared to be extended to other aspects of the portfolio. The result was that no two portfolio implementations were managed in the same way and the submissions of student work reflected the level of commitment by both teachers and students to the project. For example, one school (WA) while observed to be working on the portfolio product did not submit any portfolios.

The computer-based examination uncovered some minor manageability problems. Preparing for the examination involved a visit to discuss the task with the teacher, and on one occasion (RA) the network administrator, the pre-testing of equipment and room set up. In one school (CA) the separation between computers was minimal and students could not help but see the screens of their neighbours. In all other schools a separation of one workstation was possible. Two schools (CA, RA) were able to suspend internet access for the duration of the examination. The audio response introduced some disruption to the flow of the examination and although the noise level was low, it was not possible to know for certain if students were talking to each other or recording their reflections. Because the examination was tightly structured into sections with recommended timings and clear instructions (student forums commented favourably on the latter), it was managed in a far more standardised and consistent manner across all schools than the portfolio. Operating systems and application software available to students were nearly identical and hardware was in general up to date and matched to the demands of the task, provided that system resources were managed with a little care by not attempting to work on multiple applications.

Group and room size, as well as the layout of workstations, varied across schools. For the examination proximity of students became an issue, particularly at CA, where separation between students was only 30 cm meaning that students work was easily visible to neighbours. The opportunity for cheating was a particular concern for one teacher (WA) and there was the suggestion of design ideas *spreading* between adjacent students as described in Chapter Five. In most schools, there was ample space and spare computers were available in the event of crashes and hangs. Although disconcerting for the few students involved (at RA, WA and XA) the examination was able to be continued with minimal interruption and negligible disturbance to other students after events of this type occurred.

With the portfolio, the problem of authenticity of student work is difficult to manage since students cannot be expected to work in isolation over an extended time frame. Indeed collaboration, in the form of *show and tell* and peer review is often a necessary component of product development and was observed during the visits. The ease with which work can be taken home introduces a further dimension of uncertainty and assistance from external sites and sources cannot be discounted. The only realistic management strategy involves students and teachers signing a verification of authenticity. This was not included in this study.

Portfolio submissions presented some manageability problems with difficulties encountered with file sizes and file types. Students were frequently unaware of how big their work had become. For example at XA, one student made a safety animation using presentation software, with path animations within slides and slide transition timings set automatically to achieve the effect of motion. This file alone ran to 24 MB and presented a problem for the markers since it had to be downloaded to be viewed. Although clear instructions were given regarding the final format of files, these were often overlooked meaning that a student’s original files had to be converted before uploading to the marking repository.

In the examination, instruction as to the allowed file types was given and marks were awarded for correct implementation. Nonetheless, some files were still received in incorrect formats and had to be converted. With 4GB of storage capacity available, file size was not a consideration for students. Invigilation of the examination went smoothly in all schools with the exception of the audio recording section which disturbed the flow and added further problems to secure invigilation. Students appeared to be fully occupied with the practical section, where the number of tasks kept most working right up to the end. Students frequently mentioned the time pressure they felt during the student forums whilst maintaining that the examination was a fair and reasonable test of their abilities.

In summary, there were no *critical* manageability issues with implementation of the examination across the seven schools. However, ensuring consistency of implementation of the portfolio depended entirely on the teacher and the variability observed here points to the crucial nature of this factor. The slight variations in hardware and software between schools had little impact since what students were asked to do was developed from and closely aligned to the AIT course and, in order to offer the course, schools had to provide students with adequate and appropriate resources. Students were reasonably familiar with the applications they chose to complete the portfolio and the examination as evidenced by the sophistication and completeness of the some of the work submitted for these tasks.

Technical Facility Dimension

Technical facility in the context of this study refers to the extent to which existing technologies were suitable for adaptation to the purposes of assessment. Likely factors impacting on the technical dimension were, for example, availability of software capable of being used to develop solutions to the tasks, ability of hardware to run the students' choice of software reliably and at acceptable speeds and ease of recovery in the event of any system failure. A summary of some of the findings from observation, survey and interview data are presented in the Table 6.3. For the portfolio, there were no critical technical problems since the extended time scale allowed for the occasional delay and backing up meant work could be revisited later.

The main technical problem with the examination was sound recording where widespread technical difficulties were encountered even though this should have been problem free. Teachers and network administrators knew well in advance of this requirement and were provided with a headset for testing purposes. Students were asked to perform a sound test before commencement of reading time and equipment failure at this stage added to student anxiety. Causes were mainly due to network restrictions (RA, WA) put in place to stop students downloading music files and though these were easily removed, reinstallation of sound card drivers was required. Many students lacked the skills to make a sound recording and sometimes it was difficult to know if the fault was with the student, hardware or the software.

Choice of software impacted on processing speed particularly of image files. All schools had sophisticated multimedia software for example *Dreamweaver*, *Photoshop*, *Adobe Fireworks*, *Flash* and *Illustrator*, but often, for example at CA, RA, XA, WA, the hardware lacked the capability to run these effectively, particularly when multitasking. Occasionally during the examination computers froze (four at XA and three at RA) and had to be restarted. Where space permitted, students were moved to spare computers taking their work with them on the USB flash drive. The ease with which student work files could be transferred over highlighted a technical advantage of the use of portable memory devices. Clearly, on the technical dimension, the portfolio was more feasible than the examination mainly due to the inclusion of the audio reflection in the latter which was responsible for the majority of the problems encountered.

Table 6.3

Technical Dimension: Findings from Feasibility Framework by Case

Case	Portfolio	Examination
CA	Students enjoyed a good range of application software (Office 2007 and Adobe CS3). Hardware adequate but students commented on and perceived a need for newer and faster computers Resource hungry applications such as Photoshop caused delays and occasional hangs. .	Some technical problems: widespread failure of sound test for audio section-some down to students lack of skills. Three computers stopped responding with multiple applications open. Students not managing system.. Photoshop observed to run very slowly suggesting inadequate memory. Scans of designs hard to read due to light pencil drawings.
LA	Up to date systems with adequate memory presented few problems. Good range of application software available (Office 2007 and Adobe CS3). Student voiced concerns about potential problems with hardware but these were not observed. Delays often due to multitasking applications.	Some user errors with sound recording test caused disruption but this was before commencement. Easily fixed so audio section completed here. Student concerns about potential problems with hardware mentioned in student forum.
MA	Hardware and software (Office 2007 and Adobe CS3) easily capable of running application for all tasks. Students used Dreamweaver for website creation. Teacher experienced and *taught* application prior to commencement of task.	Teacher planned for possible problems by preparing additional workstations. Audio test completed without problems-students well prepared for this 3 USB flash drives contained corrupt files but quickly re-imaged from spare. Anxiety but minimal disruption.
RA	Wide choice of applications available (e.g. Office 2003, Adobe CS3, FrontPage); but these were server delivered and delays were observed and commented on by students and teacher. One student used open source web development software brought in from home on USB drive.	Server delivered applications led to a few hangs and restarts. Students able to move to spare machines. Sound cards had been disabled to prevent students listening to music files and were supposed to be fixed. Plenty of advanced warning and assurances from network administrator however widespread problems (continued)

		when sound recording test undertaken prior to commencement. Some students lacked know how. Sound test was eventually completed and audio section included.
WA	Good range of software (Office 2003 and Adobe CS3). Hardware up to date but teacher commented on network problems arising from server delivered applications and storage. Novell network with some rights problems preventing some students with incomplete set of application software.	Some technical problems relating to network rights and software installation caused surprise absence of some key applications from some machines. Class split and half moved to adjacent lab. Sound drivers were supposed to have been enabled but rights denied for all student users. Sound recording done separately on teacher's laptop. Teacher frustrated with problems.
XA	Abundance of application software (XP OS, Office 2007 and Adobe CS3 plus Corel Draw PhotoPaint and RAVE). Hardware differed between two labs with older machines limited to 512 MB RAM made running Adobe suite (e.g. Photoshop) sluggish with large files.	Four computers froze when opening and running Photoshop (all were 'older' computers with 512MB RAM) and had to be restarted. Two students moved to spare machines to save time. Caused anxiety but students allowed extra time as required.
ZA	Standard range of application software. XP OS, Office 2007 and Adobe CS3 Hardware up to date and effective. Server fault led to loss of backup copies by teacher. Student work eventually recovered and submitted.	No technical difficulties encountered during examination except recording of student forum at the end. This was done on researcher laptop.

Functionality Dimension

Functionality in the context of this study refers to the validity and reliability of assessor judgements made of student performances on digital forms of assessment and comparability with other methods of assessment. The Functionality dimension was divided into findings regarding validity and those regarding reliability and these are discussed in turn in this section.

Validity was analysed by considering:

i) how well the performance of students matched the curriculum outcomes;
ii) the extent to which the method of representing performance was authentic;
iii) whether or not the task and context were meaningful and relevant to students and community practice.

Great care was taken in the creation of the assessment tasks to ensure validity to the AIT course. A situation analysis (refer to Appendix B) was undertaken to make sure that the scope, difficulty and nature of the tasks matched the course outcomes and abilities of the students. The situation analysis identified those areas of the course which would be difficult if not impossible to assess using written examinations and these together with the contexts in which the tasks were to be situated are set out in the Rationale for Assessment Tasks (Appendix C). The main tasks for both the portfolio and the examination matched the overarching theme of ICT in business and were set in realistic and authentic contexts. The assessment criteria for the portfolio and the examination followed closely from the course syllabus as set out in Appendix D and Appendix E.

The comments of teachers and students with regard to the level of difficulty and appropriate nature of the tasks together with the close matching of assessment tasks to course outcomes (refer back to Table 6.1) attest to high face, content and construct validity for both the portfolio and the examination. Students' perception of the validity of the portfolio and the examination may also be inferred from responses to certain items in the questionnaire, for example, "Overall it was better doing the exam using a computer than on paper". Responses to these questions were aggregated into the scales *eAssess* (for the examination) and *eAssessP* (for the portfolio) and these ranged from 3.2 to 3.6 and 3.2 to 3.4 respectively on a scale of 1 (strongly disagree) to 4 (strongly agree).

Reliability describes the accuracy and precision of the measurement of the performance by the assessors scoring the tasks (measured by correlations between scores awarded), the extent to which teachers' scores, where available, and assessors scores were consistent and the extent to which different methods of assessment (analytic v pairs) were correlated. Table 6.4 (Portfolio) and Table 6.5 (Examination) present evidence from each of the schools in support of the assertion that the assessment tasks were valid and that the assessment marking process generated reliable scores.

In summary, in all cases students readily perceived the assessment tasks (both portfolio and examination) to be authentic and meaningful in the context of their course using words like *fair, simple to understand and easy to follow, summed up what we have done this year.* The structure of the tasks in the portfolio and the examination were open ended and designed to allow for a large range of responses. For example, in the examination, graphs of climatic data were to be included in the marketing of a resort.

Table 6.4

Portfolio: Functionality Findings from Feasibility Framework by School

Case	Functional: Validity	Functional: Reliability
CA	Experienced teacher with in depth knowledge of course requirements selected alternative task (e-commerce website) authentic though difficult many students attempted front end only. Some tried to include database to extend. Researcher and teacher agreed portfolio activities appropriate to level, based on course outline and in correct context- AIT in business.	Strong and significant correlation between scores by assessors' (r=0.87, p<0.01 paired zero scores excluded). Average of assessors' portfolio scores and Teacher semester marks moderately but significantly correlated (r=0.50, p<0.05).
LA	Experienced teacher with in depth knowledge of course requirements used tasks supplied "very similar to the task that I had intended to deliver" but assessed differently. Teacher felt tasks were appropriate to course level and allowed for student capability.	Weak correlation between external assessors' portfolio marks (r=0.48). Average assessors portfolio and Teacher semester marks were very weakly correlated (r=0.30). Small (10) sample size.
MA	Experienced teacher with in depth knowledge of course requirements used tasks supplied. Teacher followed portfolio requirements to the letter and agreed these were based on course description. Teacher felt work reflected student capability but marking guide not adequate. Portfolio formed part of semester assessment as intended.	Strong and significant correlation between external assessors portfolio scores (r=0.78, p< 0.01). Average of assessors' portfolio scores and Teacher semester marks strongly and significantly correlated (r=0.87, p<0.01). Only here was the same work marked for the portfolio by teacher and assessors.
RA	Teacher followed portfolio requirements and these were based on course description. Students positive about "doing it on the computer" although inexperienced. Teacher positive about tasks describing them as "appropriate for my cohort of students".	Strong and significant correlation between external assessors portfolio scores (r=0.90, p< 0.01 paired zero scores excluded). No teacher semester mark supplied.
WA	Experienced teacher with good knowledge of course requirements. Portfolio task (30 second cinema advert) developed by teacher with real client who visited students and provided task parameters. Teacher didn't like design process documentation components. Felt student work was poor.	No portfolios or semester marks received from teacher.
XA	Teacher developed own portfolio task- an animated safety feature broadly matching course objectives and set in teenage context. Narrow range of skills observed in digital artefacts.	Strong and significant correlation between external assessors portfolio scores (r=0.90, p<0.01 paired zero scores excluded) No teacher semester marks received.
ZA	Experienced teacher with good knowledge of subject. Used Portfolio tasks supplied as extras to own course projects-these focussed on multimedia applications. Tasks were completed though not assessed by teacher.	Strong and significant correlation between external assessors portfolio scores (r=0.79, p<0.01) Average of assessors' scores and Teacher semester marks not significantly correlated.

Table 6.5

Examination: Functionality Findings from Feasibility Framework by school. (*Key; s=Survey, t=Teacher, f=Forum, o=Observation)*

	Functional	
Case	Validity	Reliability
CA	Teacher interview- positive about tasks "students able to demonstrate many aspects of syllabus" [t]. Students found exam 'appropriate' and 'enjoyable' [f] preferring computer-based work to written work. [s]	Strong and significant correlation between external assessors' examination scores (r=0.73, p<0.01). Assessors average examination score and teacher's examination score moderately but significantly correlated (r=0.59, p<0.01).
LA	Student panel indicated that exam was a fair and appropriate assessment [f]. Teacher felt work reflected student capability [t].	Correlation between external assessors' examination scores not significant. Small (10) sample size. No teacher examination marks received.
MA	Teacher described practical task as "good" but less enthusiastic about theory [t]. Criticised marking key. Teacher felt work reflected student capability but marking guide not adequate [t]. Full compliance with all parts of study supports view that teacher perceived tasks as suitable at this level [t].	Correlation between external assessors' examination scores not significant. Moderate correlation between assessors' average examination score and teacher's examination mark (r=0.58, p<0.05).
RA	Teacher positive about examination tasks and used this as final semester examination. Teacher agreed level appropriate based on course description [t]. Students positive about "doing it on the computer" although inexperienced [s]. Students welcomed idea of computer-based examinations for other subjects [f].	Only weak correlation between external assessors' examination scores (r= 0.50, p<0.05). Correlation between assessors' average examination score and teacher's examination mark not significant.
WA	Teacher positive about examination tasks particularly practical tasks-felt the theory was excessive. Teacher reported that mark key re-written to put more emphasis on practical skills displayed. Teacher didn't like design process components in examination. Felt student work was poor [t].	Moderate but significant correlation between external assessors' examination scores (r=0.57, p<0.05). No examination marks received from teacher.
XA	Teacher felt exam was at an appropriate level for students and matched course description [t]. Student forum thought examination was "too easy"[f].	Moderate and significant correlation between external assessors' examination scores (r=0.61, p<0.01). Moderate correlation between assessors' average examination score and teacher's examination mark (r=0.66, p<0.01)
ZA	Experienced teacher with good knowledge of subject. Examination completed as extra to the course Teacher felt tasks were too simple for this level and wanted more choice and scope to demonstrate greater sophistication in skills [t]. Coursework and assessment had focussed on multimedia/ web page design [o]. Students well equipped and prepared for exam theory and practical [o].	Correlation between external assessors' examination scores not significant. No teacher examination marks received.

Though most students were able to create a simple chart of either rainfall or temperature, none were able to develop a standard climatic chart combining both, with a line graph of temperature and a column graph of rainfall and a correctly labelled and scaled axis for each. It was therefore possible to award a wide spread of scores making possible discrimination between candidates of widely varying abilities.

For the portfolio, correlations between the scores awarded by the two external assessors were strong and significant in all cases with the exception of LA where the sample size was small. For the whole sample of 115 students, the *total* scores awarded by each external assessor were very strongly correlated ($r=0.89$, $p<0.01$). The Cronbach Alpha statistic will generally increase as the inter-correlations among test items increase and hence it may be used here as a measure of the internal consistency or reliability of the test scores awarded. Rasch analysis of the portfolio scores using a polytomous model generated a Cronbach Alpha coefficient of 0.94 ($N=115$) that indicates a high internal consistency between assessors. (Note that a reliability coefficient of 0.7 or higher is considered acceptable). For the three components of the portfolio, the reliability coefficients between assessors for each component of the portfolio were 0.92, 0.94 and 0.87 respectively.

For the examination, there was much less consistency and correlations between assessors scores varied between strong and significant at CA ($r=0.73$, $p<0.01$) to very weak and not significant at ZA ($r=0.23$) with most being at best moderate and significant at the 0.05 confidence level. Possible explanations for these variations might include marker error, errors in interpretation of the analytical marking rubric or faults with the rubric itself making accurate scoring difficult. For the whole sample of 115 students, the total scores awarded for the examination by each of the two external assessors were moderately correlated with each other ($r=0.61$, $p<0.01$).

Marking of the portfolio and the examination were achieved with acceptable reliability given that assessors did not meet to discuss interpretations of the marking rubric, or practice, or engage in any attempt at standardisation or moderation. Discrepancies in scores awarded were not examined and there was no attempt at re-marking. It is likely that even greater accuracy could have been achieved had standard practices like those mentioned above been adopted.

Pedagogy Dimension

Pedagogy in the context of this study refers to the extent to which the digital form of assessment supported and enhanced teaching and learning; in other words the extent to which assessment was aligned with classroom practice. Discussion with teachers and students and observation of classes in action allowed a composite of principles and methods to be described. Teachers involved in the study were all passionate about the AIT course and the opportunity afforded to all students, particularly those who might not excel in traditional school subjects and settings. They all recognised the links between AIT skills and real world problems and the fact that these are often complex and open ended with various possible solutions. The implicit and fundamental underpinning of the technology process and the creative application of information design principles was also well recognised and practiced. The portfolio task was designed to form part of the semester's work being aligned as closely as possible with preferred pedagogy. The examination was developed directly from the context of the AIT Stage 2 course which includes the following:

> …application/use of common ICT business software including descriptions, examples and use of: personal information managers…presentation software for business… word processing simple spreadsheets basic formulas and charting…flat file databases… business Office Suites…online office applications…publishing. (Curriculum Council of WA, 2009)

Not all of these aspects could be included in the two hour examination but as many as possible were included. Table 6.6 summarises the various aspects pertaining to the pedagogy dimension. The source of each item is denoted by the code letter following the statement.

Typically students liked the idea of a practical assessment task in preference to a written paper using words like *much better* and *enjoyable* in the comparison. Teachers overwhelmingly felt that computer-based assessments matched the intended pedagogy for the course and were often complimentary about the tasks themselves as indicated by comments in Chapter Five. Most students believed they could demonstrate their full ability on the computer and that computer work was an essential component of demonstrating their ability in AIT as evidenced by results of the student survey. Students were very positively disposed to completing these types of assessments on computers. Students responded positively to statements describing the *ease* of assessment by computer-based examination (*eAssess* mean=3.20) and digital portfolio (*eAssessP* mean=3.16) on a scale of 1 (strongly disagree) to 4 (strongly agree).

Table 6.6

Pedagogy Findings from Feasibility Framework by School. Key; s=Survey, t=Teacher, f=Forum, o=Observation

Case	Pedagogy	
	Portfolio	Examination
CA	Teacher and students indicated that assessment matched typical classroom practice [t, o]. Design process document didn't exactly match teacher requirements [t].	Skills required were typical of those acquired and used in class activities [t, f]. Provided opportunity to practice and extend skills [t].
LA	Teacher and students indicated that assessment matched typical classroom practice [t, o]. Teacher not sure how much of assessment to allow at home [t].	Skills required were typical of those acquired and used in class activities [o]. Provided opportunity to practice and extend skills [o]. Teacher strongly in favour of practical component [t].
MA	Teacher and students positive towards portfolio which matched typical classroom practice and assessment [t, f, o]. Teacher indicated difficulty addressing all content of syllabus [t].	Skills required were typical of those acquired and used in class activities [t, o]. Provided opportunity to practice and extend skills [o].
RA	Teacher and students indicated that assessment matched typical pedagogy [t, f].	Skills required were typical of those acquired and used in class activities [t, o]. Provided opportunity to practice and extend skills. Students strongly preferred practical work over theory [s].
WA	Teacher very positive about making assessment structure match class work [t].	Skills required were typical of those acquired and used in class activities [t, o]. Provided opportunity to practice and extend skills. Students liked practical exam but not reflective questions [f].
XA	Portfolio not included as part of school-based assessment [t].	Activity matched classroom practice [o]. Skills required were typical of those acquired and used in class activities [o]. Provided opportunity to practice and extend skills [f].
ZA	Portfolio matched and extended typical classroom practice [s, t] but done as an extra activity and not assessed [t].	Skills required were typical of those acquired and used in class activities [s, o]. Provided opportunity to practice and extend skills [f].

Though students indicated that they had little experience in doing examinations on computers (44% indicated no experience) 62% felt they would need little or no time to get used to the process. Most of the rest felt they would need some time to do so. Almost all students indicated that doing the examination on the computer was quick, easy and preferable to the traditional pen and paper examination. Only for development of design ideas did preference fall below 70%, reflecting the observed preference of students to complete designs on paper.

Reliability of Assessment by Multiple Comparisons of Pairs

The fourth subsidiary question addressed the reliability of assessment by multiple comparisons of pairs. Only the practical component (component 5) of 60 students was selected for marking by this method. Unique sequences of comparisons were made by each of five assessors, two of whom were the analytical markers, and these in turn led to a ranking of each students work and interval scores resulting from the application of a Rasch model. For the comparative-pairs marking, the inter-rater reliability may be measured by the Separation Index statistic; the Cronbach Alpha statistic cannot be used because there are no numerical scores to be compared. The Separation Index should be between 0 and 1 (as described in Table 4.10 of Chapter Four) with values close to 1 indicating good internal consistency or reliability. A separation index of 0.96 for the holistic judgments of the five assessors indicates high reliability.

The fifth and final subsidiary question concerned the relative merits of each method of marking. There was a strong and significant correlation ($r=0.73$, $p<0.01$) between the score generated by the comparative-pairs marking (Rasch location value) and the score produced by averaging the scores awarded by the two assessors using the analytical marking rubric for the practical component. There was no significant correlation between the teacher's examination score and the comparative-pairs score. The correlation between the marking methods, though strong and significant, might be expected to be stronger and the reasons for the discrepancy required further investigation. Looking at the ranking of each work sample by each marking method showed that the greatest anomalies occurred in work which was inconsistent, that is, good in some parts and poor in others. Tables 4.16, 4.17 and 4.18 in Chapter Four examine the ten greatest anomalies. It is in these cases that differences in the marking criteria and weightings had the most effect often producing a wide discrepancy in the rank order of the work. Where work was of a consistent standard, differences in emphasis of the marking criteria appeared to have less of an effect. What is required is for the marking criteria to be made as similar as possible without regard to the marking method (analytical or comparative). In this study, this was not the case and although a strong correlation is exhibited between methods, this would surely have been even stronger if judgements had been based on identical criteria. The marking criteria for the analytical and pairs marking are set out in Appendixes E and L. It must however be accepted that the different methods can and do produce some variation in rank order and hence the grade awarded to a piece of work may vary depending on the marking method chosen. Since the variation is across the board this might well produce differences in grade boundaries and the pass/fail cut off dependent on the method of marking.

The choice of method probably depends on the extent to which the assessment task is holistic in nature. For tasks where an overall impression or impact are valued, as was the case with the brochure and the marking criteria in this study, comparative-pairs assessment proved to be a quick and reliable method of discriminating between student work samples. However, due to the limited nature of this trial more research is required and this was done in later phases of the main project.

Constraints and Benefits

From the case studies, it was possible to draw out a summary of the main constraints and benefits of assessment by digital portfolio and assessment by computer-based examination. These are shown by case in Table 6.7. Both the portfolio and the examination had different constraints. For the portfolio these centred on manageability. The requirements of the portfolio and imposition of the study on lesson time were not fully appreciated by some of the teachers. The external assessment requirements of the course naturally took preference and preparing students for an external theory examination took up time which might have been devoted to the portfolio and comments to this effect were expressed by teachers at LA and ZA. This limited the quality and degree of completeness of many of the portfolios submitted. The requirement that a student’s work be entirely their own was also impossible to check or to manage. Collaborative learning approaches and group work were commonly observed and it would have been unrealistic to have students work in isolation. Further, with portable digital storage and email, students could easily work on their portfolios at home, effectively unsupervised.

Constraints on assessment by computer-based examination centred mainly on technical issues. Although there were no critical problems, there were quite a few minor technical glitches which though easily and quickly overcome, did add to the anxiety experienced by students and invigilators alike.

Table 6.7

Assessment by Portfolio and Examination: Summary of Constraints and Benefits by School

Case	Constraints	Benefits
CA	All students able to complete portfolio and examination using hardware and software supplied. Time allocation sufficient though students suggested extra time for practical component of exam. Some students "not looking forward" to exam according to teacher. A few technical glitches.	Engagement/positive student response and opportunity to demonstrate creative solutions and practical skills. Students admitted to writing more when keyboarding responses to theory section of exam-easier and more accurate.
LA	All students able to complete portfolio and examination using hardware and software supplied. Time allocation for exam sufficient.	Motivated students felt tasks were 'worthy' of additional effort. Positive attitude towards assessment using computer. Students enjoyed doing examination and preferred keyboarding to writing. Spell and grammar check. Quantity and presentation superior to handwriting.
MA	A few technical glitches easily fixed.	Positive student response to examination – easier, faster, more accurate and creative. Fixing errors, quantity and presentation of work all perceived as benefit by students.
RA	Language used in portfolio instruction and design document questions challenged comprehension of some students. Examination had some technical difficulties due to network set up. Nothing critical but disconcerting for the few students affected.	Positive students' response to computer assessment. Exam fully engaged class. Students said that computer-based exam was "much better". Some students engaged with portfolio-creative solutions.
WA	Portfolio product became too complex for some students who opted to use advanced software and then became bogged down. Unforeseen technical problems in examination caused disruption.	Students very positive about assessment with a practical bias.
XA	Language in reflective questions challenged students understanding. Portfolio was not managed according to required parameters.	Teacher and students all preferred practical exam to theory paper. Portfolio gave scope for student skills and creativity.
ZA	None evident other than time allocation due to non-inclusion with school-based assessment.	Students' positive towards computer-based assessment. Teacher also but with qualifications and suggestions for improvement.

These technical difficulties occurred in almost every school with the audio recording causing most and probably adding little information to the study. Further, the audio reflection disturbed the peace and quiet of the examination and introduced supervision problems. Catastrophic events such as a power failure would obviously have prevented completion of the examination and short of emergency power supplies or the use of battery powered laptops there appears to be no work around in the unlikely event of this occurring.

The nature of the tasks required fairly basic software and only light processing so that students using older machines would not be disadvantaged. However, the equity between schools with regard to resources might become a constraint in the event of more advanced assessment tasks being created. In interviews and surveys, students often lamented the lack of modernity of their computers and the perceived negative impact of this on performance during the examination (for example class CA). However the same students were observed to adopt poor working practices with multiple files open concurrently. Computers with more memory and multiple processors certainly would alleviate the effects of poor system resource management such as this, and in this regard could be seen as a constraint on the equity of the examination across different schools.

Some students and one teacher (ZA) were not happy with the rigid structure of the examination and suggested an unconstrained time format, more choice and more creative freedom. A further constraint of the examination was the limited scope of assessment which was possible in the short time frame.

Both the portfolio and the examination had different benefits. For the portfolio these were the greater authenticity of the tasks (e.g. class WA worked with a real business to develop advertising), the validity of the tasks in terms of their fidelity to and extent of coverage of course learning outcomes and the greater alignment of the tasks with the methods and principles of teaching prevalent in the classroom. The examination had the major advantage of being easily manageable within a typical school and valid with regard to the fact that the work produced by the students was entirely their own. Though collaboration is valued and encouraged in the world outside the classroom, the current assessment system requires that students' ability be assessed in isolation on identical tasks.

Summary

This chapter set out to address the overarching research question through firstly in turn discussing the subsidiary questions in the light of the findings of the study. The key points of this discussion are now summarised directly in terms of the research question to lead into a presentation of conclusions in the final chapter.

How may the digital capture of students' performance most effectively support summative assessment in the senior secondary Applied Information Technology course?

The central concept to the question is the capture of student performance in digital form. In general terms the performance to be captured in AIT was the student's response to a challenge given in the form of a design brief with the design and development of a prototype digital product. Thus the end product is necessarily captured in digital form. However the design and development processes also needed to be captured digitally with either the student creating these in digital form or with filming, photographing or scanning involved.

The digital capture was for summative assessment purposes. The study considered two types of summative assessment, a digital portfolio and a computer-based examination, with the main differences being the limited time, challenge and set of processes possible in the computer-based exam. In the practical component of the examination, students only had two hours, had no opportunity to investigate the challenge, were guided through the design and development processes and were more limited in access to digitising tools with, for example, no access to cameras and the Internet, while having access to the same array of locally stored software.

The study found that for both the portfolio and the examination it was possible in normal schools to digitally capture a substantial and critical component of student performance in AIT. Both forms of assessment allowed for a range of levels of performance as evidenced by the spread of marks awarded with even the more complete solutions having potential for greater development and sophistication. This was also evidenced by the stated perceptions of students, teachers and assessors.

The study used a four-dimensioned Feasibility Framework to investigate the effectiveness of each form of assessment; portfolio and examination. In terms of manageability the examination was found to be more effective due to the difficulty in consistently implementing the portfolio, with *reasonable* compliance occurring only for two of the seven schools. In particular, teachers in the others did not adequately invigilate the time allowed, the recording of processes, adherence to file format limitations and acknowledgement of assistance.

In terms of the technical dimension, though both forms were not without problems, the extended time frame allocated to the portfolio meant that the impact of any technical issues was far less significant and in this regard the portfolio was preferable. During the examination there were widespread problems with the audio recording section and with a small number of the USB flash memory modules and the fact that these had to be resolved immediately added measurably to the sense of anxiety and tension already being experienced by the students.

With respect to the functionality dimension, both formats were considered to be valid forms of assessment by teachers and students alike. The close matching of tasks to course outcomes was noted by teachers. The extended skill set demanded of the portfolio meant that its content validity was higher than that of the examination which was more limited in its extent. The marks awarded by teachers for the whole semester and those awarded by assessors for the portfolio and examination were not indicative of any predictive validity of the tasks with the exception of the two schools where the implementation was most closely aligned to that intended. The consistency of scores awarded by assessors points to high inter-rater reliability and supports the design of the marking rubrics.

In terms of the pedagogy dimension, the portfolio, with its extended scope and time frame, was more closely matched to everyday classroom practice with unrestricted access to resources and the opportunity for collaborative learning. Most teachers and students were not familiar with computer-based examinations.

In conclusion, though digital capture of students' performance is not without problems in either form, for each, the benefits far outweighed the constraints when compared with the current pen and paper based response examination. Digital capture allowed authentic practises to be to be assessed; authentic in terms of their fidelity to the course context and descriptors and authentic in terms of the real world use of computer technology. These issues will be discussed in more detail in the next chapter. The following, and final, chapter will draw conclusions from the study related to the research question and go on to recommend implications of these conclusions for practice and future research while acknowledging inherent limitations of the study.

CHAPTER SEVEN

CONCLUSIONS

This chapter brings together the main conclusions from the study in relation to the research question and goes on to recommend implications of these conclusions for current and future practice and for further research whilst acknowledging the inherent limitations of the study.

In essence the results of this study allow conclusions to be drawn with respect to the relative merits of *three* forms of assessment: the traditional pen and paper response examination, the computer-based performance task examination and the digital portfolio. The case will be made for replacement of the current traditional assessment practice in AIT with each of the two alternative forms, pointing out the constraints and benefits derived from the study. Finally, a comparison between the two digital forms will be made with recommendations for future practice and further research.

The research question for this study was:
How may the digital capture of students' performance most effectively support summative assessment in the senior secondary Applied Information Technology course?

The four key concepts within the question are *digital capture, students' performance, effective support,* and *summative assessment.* Conclusions relating to each of these aspects will now be presented. The five subsidiary research questions were addressed in the previous chapter and thus the following discussion draws on those findings.

Digital Capture

The term *digital capture* describes the creation and storage of information in binary form, for example, digital video, analogue video converted to digital form, digital audio, analogue audio converted to digital form, digital images and graphics, analogue images and graphics converted to digital form, word processed documents, spreadsheets, databases, digital slideshows and web pages. These binary forms are created and accessed using computers or computer-based technologies.

In the portfolio, all of these forms might have been employed given that the suggested design brief was the creation of an e-commerce website with the opportunity to add two further digital artefacts of the student's choice. In reality, a typical portfolio made use of some form of graphic design application and web page creation software together with word processing and presentation tools from an office package. Students were not observed to use analogue to digital devices such as cameras or graphics tablets and appeared to create and revise their work entirely on computer. Students work was collated and copied to disks for submission.

In the examination, students were provided with digital resources in the form of images and text files on a flash memory device and only these were permitted to be used in the solution. Typically, students used a graphic design package to sample, manipulate and edit the images and *MS Excel* to present the data from the text file in graphical form. Almost all used a *MS Publisher* template to scaffold the creation of their marketing brochure, create the text, and set the fonts, colours and layout adding in other resources as required with a handful using *MS Word.* In design and creation of the logo, the intention was for students to use a graphical design application of their choice to create from scratch an emblem in keeping with the nature of the task as set out the marketing brief. It was here that students had the opportunity to put into practice the design principles and conventions detailed in the course. However, most students appeared to misunderstand the concept of a logo or lacked the skills to develop one, and submissions were in general poor. The only analogue aspect of the examination was the pencil and paper used by an overwhelming majority of students in initial design sketches and these were later digitised using a scanner. All other work was saved back to the flash memory device and these were collected at the end of the examination.

The range of file types received from both portfolio and examination was limited and for many of these it was possible to convert to a common standard. For example, initial designs, *MS Word* documents and *MS Publisher* files, were all easily converted to portable document format (PDF) reducing file size and allowing for rapid display in a web browser.

The study demonstrated overwhelming advantages of digital capture in AIT when compared to traditional assessment methods. Digital capture allowed students to be assessed on a greater range of skills which comprise the technology process, the central tenet underpinning the rationale for the AIT course. In its absence, the assessment process is simply incapable of the authenticity, rigour and comprehensiveness which are self-evident requirements in the course description, for example in the requirement that students should "implement...production processes...select and use computer hardware and software...(and) apply skills, techniques and processes to develop information solutions" (Curriculum Council of Western Australia, 2009, p. 3). In this study, not only were students able to demonstrate digital skills but were also able to demonstrate creativity in design and production, planning and evaluation.

Solutions to the technology challenges or design briefs developed in the contexts of the course are by definition digital in form. Although their specific composition may vary, their existence requires the manipulation of digital content using a range of software of greater or lesser sophistication. What is created, in its fundamental nature, are simply ones and zeros which have no meaning outside the digital domain. Further, digital capture affords not just the opportunity to experience the final product but also a simple way of recording, communicating and reflecting on the creative process as it actually happens. Scans of initial design ideas, photographs of prototypes, audio commentaries and reflections, digital journals, blogs, tweets and podcasts may quickly capture and share the technology process in action in forms which are rich, varied and greatly superior to the one dimensional pencil and paper records which today's students, as evidenced by comments from the study, are reluctant to countenance and see as belonging to a bygone age.

Digital capture also affords great advantages in storage and transmission of student artefacts. Recent years have seen a huge expansion in the types and locations of digital storage. USB flash memory, employed in this study to deliver resources and collect students' work in the examination, is now so common and cheap that it is worth remembering that capacities of 1 GB were only developed in early 2005 yet by the end of 2009 a 256 GB Flash Drive had been produced. Prices of this and other forms of storage have fallen so dramatically in recent times that there is an overwhelming financial case for storing and transporting student work in digital form even when it comprises the text and graphics developed for print media. In this study, disks and flash storage were used for transport of materials although it was envisaged that these could be replaced with on-line systems.

Online digital storage has also expanded in capacity as a result of developments in non-volatile magnetic storage. With free services offering several Gigabytes of space, cost is no longer a barrier to hosted storage. Further, backing up copies of students' work in digital form is quick and cheap. Copies of work can be easily shared and transmitted by file transfer protocols even when file sizes are large, affording the advantages and economies concomitant to distributed assessment systems of the type trialled here. In this study, on-line systems and tools were successfully used to give assessors access to students' work.

The study showed that even the physical collection of student work on USB memory was workable though perhaps a little cumbersome. The devices, once used, were simple and quick to re-image and could be re-used in subsequent years. Delivering not only files and resources for the examination but also the question paper itself, led to savings in printing, security and transportation costs. It is not difficult to imagine evolution to web based storage with students downloading digital resources and questions at the beginning and uploading digital products and answers at the conclusion of an examination. Such a system has the potential to deliver huge cost and efficiency benefits.

Student Performance

In general terms, a performance comprises an event in which a person or a group of people behave in a particular way for another person or group. In the context of AIT, performance is what students do in pursuit of solutions to tasks set by teachers and examiners and includes not only the final product but also evidence of the creative method undertaken along the way, best described as the *technology process*.

The course description for AIT states that it should "provide opportunities for students to develop knowledge and skills relevant to the use of ICT to meet everyday challenges" (Curriculum Council of Western Australia, 2009, P.5). making the inclusion of performance based tasks essential. The assessment structure mandates that at least half of the assessments should be on production tasks, however the current external examination is a three hour paper based examination. Clearly then, students, teachers and the course authors expect performance based assessment in some shape or form. Tasks set in context, like the ones trialled in this study, are capable of allowing students to display a wide range of competencies, for they imitate to some degree the problems faced in life itself, which are frequently open-ended and complex. Their predictive validity, for success in the world outside the classroom, should therefore be high.

Comments from students and teachers as well as responses to questionnaires indicated that for the participants of this study the expectation, and preference, was overwhelmingly in favour of performance based assessment in some form. Students' response to questions on the ease of assessment by portfolio and by examination, represented by scale scores, revealed very positive sentiments and negligible difference between the two forms. Most students indicated that the portfolio and the practical examination provided ample opportunity to demonstrate their practical performance skills. Almost all students indicated a preference for the assessment of practical performance as a measure of their ability in AIT. They commented on the ease of working on a computer in comparison to working on paper, citing speed of action, physical comfort and ease of correcting as some of the main advantages.
Students perceived the assessment tasks to be fair tests of their ability and recognised that a course centred on the technology process must incorporate assessment of the practical skills development through it. These sentiments were echoed by the teachers in their responses to the e mail questionnaire.

Summative Assessment

The purpose of summative assessment in AIT is to identify the relative competence achieved by students in all aspects of the course. Students taking the course must proceed to an external assessment, the results of which are used to moderate scores from school based assessments and ultimately to contribute a score towards ranking for tertiary entrance. Assessment in AIT is of a high stakes nature since the results obtained may be used as prerequisites for admission to further study.

The Curriculum Framework of Western Australia (1998) set out the criteria for assessment. Summative assessment should be *valid* with judgements "based on assessment information about the outcome in its fullest sense, rather than only on some parts of it, a proxy for it or a rote manifestation of it". It should be educative and "contribute to the achievement of the learning outcomes". Further, it should be explicit with "clear and public criteria", fair, allowing reliable judgements to be made and for students to have equal opportunities to demonstrate achievement, comprehensive, with judgements of student progress "based on multiple kinds and sources of evidence. Information collected...should provide a reliable indication of whether students can do the things described in the outcomes consistently and autonomously over a range of circumstances".

If assessment tasks in AIT are to be valid, authentic and pedagogically sound they should, as argued above, be performance based and of necessity digitally captured. However, an assessment task is of little value if it is unmanageable, cannot be assessed reliably or if the cost of administration and scoring is prohibitive. Summative judgment of performance tasks is a highly skilled human activity that cannot at the present time be replaced by any mechanistic or computer-based method. This study has shown that a digital reflective process portfolio and a performance/production examination can be used for the purposes of summative assessment to achieve acceptable reliability using conventional, rubric based marking methods and a comparative-pairs marking method for the examination.

Analytical Marking

The view of the assessors was that analytical marking of the performance based tasks was no more difficult than scoring traditional written responses and though time consuming at the outset was no more onerous or cumbersome. The reliability of the summative assessment by analytical marking was high for both the portfolio and the examination and could have been higher if consensus marking had been undertaken. The markers worked entirely independently, did not discuss the rubric or its interpretation, underwent no moderation and yet still produced good agreement on scores.

The analysis of the marking, with strong correlations between scores awarded by the two assessors, pointed to the reliability of the method. After more in depth analysis, it was found that the correlation between markers for the analytical marking was more acceptable for the marking of the Portfolios ($r=0.9$, $p<0.01$) than for the Examination ($r=0.6$, $p<0.01$). Rasch polytomous model analysis yielded Separation Index values of around 0.95 for the portfolio and 0.85 for the examination. The analytical marking of these performance tasks therefore meets the measurement requirement of score reliability.

Comparative-Pairs Marking

A second method of scoring, by comparison of pairs of work samples, was undertaken by five assessors for the practical examination component (a marketing brochure for a resort hotel) of 60 selected students. In this method, assessors made a choice of the better of the two student work samples, based on three criteria before deciding on an overall or holistic *winner*. Each work sample was then involved in several subsequent comparisons until a scale of acceptable accuracy had been generated. At the top of the scale was the work sample deemed to be superior to all of the others, even though it was not in fact compared to all others, its superiority was inferred from the comparisons made with other good exemplars. At the bottom of the scale was the work sample which had been judged least favourably in its sequence of comparisons, again by inference based on the comparisons actually made. The intermediate positions were occupied by work judged to be of increasing merit.

This method of marking was found to be highly reliable in discriminating between students on performance based tasks. The reliability of the results of marking was high with a Separation Index above 0.9 for the practical component of the examination. The reliability comes from the elimination of marker bias; it matters not how many marks are awarded but merely which of two samples was perceived to be superior. Strict or generous markers, who might produce widely varying scores when applying a marking rubric to the same piece of work, tended to value the *relative* merits of two pieces of work in the same way leading to high inter-rater reliability.

The time taken to assess by *comparative-pairs* was found to be initially a little longer than with the analytic marking rubric, particularly where the work samples were closely matched. However, with practice and use of the comment field associated with each sample, assessment quickly became quite rapid. In many instances it was obvious, at a glance, which of the two work samples was the better. The assessment system used predetermined sets of comparisons, not comparisons generated by the most recent comparison and this threw up quite a few obvious mismatches. However the system, even in this rudimentary form was capable of scoring performance in digital form for summative assessment with very high reliability. Once again the assessors did not meet or practice prior to the assessment and although they were five in number, the time taken overall was not reported to be onerous. On the contrary, the simplicity and speed of this system was perceived by some as superior to the analytical marking of the same work samples. There were differences in the rankings derived from the two methods of marking (analytical v pairs) but overall the correlation between the two was strong ($r \approx 0.73$, $p<0.01$).

Conclusions about the Assessment Tasks

All seven classes in the study attempted the five components of the assessment task to varying degrees of completion. Three components were part of a digital portfolio; the other two were parts of a computer-based performance focussed examination.

The portfolio, comprising a product, process document and two further digital artefacts, was intended to form part of the school based assessment structure for the course. For the small group of students from the two schools where the portfolio was implemented more or less in line with the intention of the study, the portfolio worked well. However, there was misunderstanding of requirements among the teachers and five of the seven ran the portfolio in whole or in part as an additional task, not counting towards the student's final semester mark. It is therefore not surprising that many students did not give the portfolio their best efforts and many portfolio submissions were incomplete. The concept of the portfolio product was well understood but some confusion was evident over the intention of the process document in spite of the detailed explanation of the structure supplied with the task specification. The digital artefacts submitted were in general disappointing. Students rarely took the opportunity to showcase a broad skill base or sophisticated technical proficiency. Frequently the artefacts were of the same type or missing altogether.

The three-hour examination paper, consisting of a one-hour keyboarded response and a two hour practical test, was attempted by all students. The typed response section consisted of several reflective questions about the portfolio task. Students' responses suggested that either the intention of the questions was unclear or that they did not know how to answer. One student commented "The wording in the second part…a bit confusing. I had to guess at what it meant", and another, "It took me a while to work out what was meant by it". There was widespread confusion over the stages of the technology process and the distinction between these with many responses repeated. As one student noted, "it just seemed like you asked the same questions four times…I got four words out of a thesaurus and copied and pasted those in three or four times".

From observation, marking and comments made by students, the intention of the practical component of the examination was clearly understood by all participants, namely to design and create a logo for a business and incorporate that into a tri-fold advertising brochure. However, there was widespread variation in students' interpretation of what constituted a logo. More than half the students simply added a caption to one of the photographs supplied. Less than 20 students designed a logo and used drawing tools to create it and even here, many used themes or cues from the photographs supplied. It was as if the photographs closed off creative avenues and confined students' imagination. Only five students did not use a software template for the tri-fold brochure and of these four used a word- processed document with three columns. In general, students selected appropriate photographs for the brochure showing that they understood that the intention of the task was to market a resort as luxurious yet having a low environmental impact. Student audio reflections, where available, supported this understanding.

In conclusion, there is evidence that the practical examination worked well as an assessment task for AIT but allowed only a relatively narrow range performance to be demonstrated. This was a consequence of the two hour time allocation and is evidenced by the narrow range of scores awarded under the analytical marking rubric. Whatever the context of the examination, the restricted time allocation and reduced scope of potential assessment tasks, together with the limited skill set required to complete them remains the major weakness of this method of assessment.

Effective Support

Digital technology in education is an all-embracing term for the ever-evolving computer hardware and software used in teaching and learning in schools, at home and beyond the home and school environments. Digital technologies employed in the portfolio were typically a combination of desktop computers, the internet, an office application suite and a graphical design application suite. During school visits, the use of peripheral devices such as digital video and photographic cameras was not observed. Writable disks were used to collect student work. In the examination, the same software was employed and internet access was disabled in two schools and not observed to be used in the others. Portable 4GB flash memory devices were used to deliver and collect student work.

Digital technologies underpinned the assessment process detailed in this study from beginning to end. Without digital technologies assessment of students' performance in AIT would not have been possible with the exception of those processes, such as design sketches and design documentation, which are amenable to development in alternative forms. A typical student's performance included creation of a multi-page, static HTML website with an animated banner, roll-over buttons, menus, graphics, text and hyperlinks, and a tri-fold brochure including a custom designed logo, manipulated images, graphs, tables, charts and text all created and stored as digital files. Digital technologies not only supported but enhanced the assessment process, allowing new techniques and methods to be developed, for example, the creation of the web-based repository of students work, the development of the on-line marking tools and the delivery of selected resources and data for the examination.

In spite of a few technical obstacles, performance-based assessments of the type investigated here, supporting the richer and more open-ended curriculum that the authors of the AIT course surely intended, were found to be a feasible and manageable alternative to traditional forms. Digital technologies allowed for more valid and accurate assessment of the diverse skills of the students; skills that would have remained unrecognized under the present paper based system. Such skills included creation of simple animations, manipulation of digital images, creation of roll over buttons, menus and hyperlinks, creation of tables, graphs and charts, involving a range of software application techniques and file types. Performance assessment was shown to be easily manageable using only the everyday technology found in all the participating schools. Further, digital technologies offer scope for increasingly sophisticated methods of teaching, learning and assessment which could further refine the methods of assessment described in this study. Information technology support for summative assessment will now be discussed in terms of the methods of marking and forms of assessment tasks.

Conclusions about Methods of Marking

The collation of student work in digital form had several obvious advantages for task assessment such as ease of storage, backup, transmission, access and sharing. The on-line database, which held the student work, was responsive and easy to use provided that adequate Internet bandwidth was available to the marker. Marking was possible, and indeed took place, from countries outside Australia. The opening of large files from remote locations sometimes presented delays when internet speeds were slow and on occasions the marking system appeared to stop responding. These delays were not experienced from connections within Australia. The marking system was never off- line and always quick to respond to marker input.

Two methods of scoring students work were used; an analytical method of summation of marks for each part based on a marking rubric and a method of comparative pairs. With regard to the analytic marking, the ability to view both the work sample with the marking rubric alongside it was convenient and ensured focus was maintained. Switching rapidly between different aspects of student work was easy. The database recorded and summed the scores and this was obviously quick and accurate. After a little practice with the system and a brief familiarisation with the criteria, the comparative pairs' assessment was also quick and convenient. For many of the comparisons, it was immediately obvious which piece of work was superior. Only occasionally were the two samples so close that it took time to arrive at a decision. Overall, the use of online technologies and digital repositories supported the marking of both forms of assessment and both methods of marking.

Conclusions about Forms of Assessment

Three forms of assessment were investigated in this study. The current system, a three hour paper based examination, and two computer-based systems, a digital portfolio and a computer-based examination. Comparison will first be made between the current system and each of the two computer-based systems. Subsequently, each of the two computer-based systems will be compared and contrasted.

The Case for a Digital Portfolio

The current AIT syllabus states that the intention of the course is to provide "opportunities for students to develop knowledge and skills relevant to the use of ICT to meet everyday challenges". It follows that during the course students should "consider a variety of computer applications for use in their own lives, business and the wider community" (Curriculum Council of Western Australia, 2009, p.3). In everyday practice students spend the majority of their time in class using digital technologies to develop information solutions and yet the current form of external assessment consists solely of a three-hour paper-based examination. This is despite the fact that the syllabus stipulates that between 50-60% of the weighting of assessment should be on production tasks. Clearly the intention of the course is to be product focussed and the current external assessment does not recognise this. Replacing the external paper based examination with a student digital portfolio would allow a realignment of assessment practices to more closely match the intentions of the course.

Almost all students indicated a preference for the assessment of practical performance using a computer and most indicated that a digital portfolio provided a fair and just assessment of practical performance allowing them to demonstrate their capability. Students commented on the ease of working on the computer compared to working on paper citing examples such as correcting errors, speed of writing, amount of writing, speed of action and physical comfort. However, the manner in which the portfolio was implemented across the seven schools in the study highlighted some areas of concern which continue to adversely influence the chances of its adoption as a high stakes summative assessment instrument.

Principal amongst these concerns was the lack of consistency in approach taken by teachers. Although some freedom was given in interpretation and task selection and although all teachers attempted to adhere to the requirements, the variations in implementation call into question the fairness of the assessment for high-stakes purposes. Success in this endeavour tended to be connected with the extent to which the portfolio was included as a part of the school-based assessment rather than as an additional task. In general, students provided a more complete portfolio where it was included as part of their school-based assessment and it is self-evident that greater attention would be devoted to a portfolio that really did form part of an external assessment process.

Requirements of component one, the product of the portfolio, were well understood and provided adequate scope for students to demonstrate their capability. For component two, the process document, requirements appeared to be less well understood and submissions varied considerably in quality with some lack of understanding of the technology process. Students needed to select the best information to include within page limits and many appeared to have difficulty with this. Component three, the two extra digital artefacts provided scope for the presentation of a broader range of skills. Unfortunately many students did not make use of this with the submission of two similar examples of their work sometimes using the same application. Even in the better submissions, the majority of students didn't demonstrate a breadth of skill. However, some students did and included the half-page information sheet explaining their artefacts.

The results of marking using the analytical rubric-based approach provided a good spread of scores for the portfolio with very high overall correlations between the two markers (around $r = 0.9$, $p<0.01$). Some variation between markers was noted in the individual components and the initial high correlation between scores for the digital artefacts was found to be largely a result of many of these being missing and scored as zero. Even allowing for this, the correlations were in general good (around $r=0.6$, $p<0.01$). There were also significant, moderate correlations between the scores on the portfolio and examination (around $r = 0.5$, $p<0.01$). There was little correlation between any of these marks and those provided by the teachers except for the two schools that most rigorously implemented the portfolio. Most teachers did not provide a set of marks for the assessment task but rather a semester mark and grade.

Notwithstanding the above, the portfolio was found to be a feasible, valid and pedagogically sound alternative assessment form. In two of the seven cases, it was demonstrated that it could be implemented effectively. However, a major mitigating factor, overhanging its potential as a replacement, centres on the originality and authenticity of students' work. The AIT syllabus rationale describes the requirement to work "both independently and collaboratively" (Curriculum Council of Western Australia, 2009, p. 3) and whilst this intention is laudable and realistic, it undermines the use of the current assessment system as a selection instrument where it is the performance of the individual which is currently measured in isolation from the people, tools and technologies which make up the education process. In the present climate, there would be little public confidence in a system of selection in which collusion between candidates or reliance on outside help might occur. With a school based portfolio, and in some cases the opportunity to take work home, this would be difficult to take into account. Paper-based assessment, under strict examination conditions, does not present these problems and currently remains the preferred, high-stakes method of judging student ability.

The Case for a Computer-Based Performance Task Examination

This section makes comparison between the current pen and paper assessment practice in AIT and a computer-based examination including a performance task. The case against a solely paper based examination rests on the contradictions arising from the AIT course outline which includes explicit statement of the practical intention of the course and stipulation as to the weighting of practical/ performance assessment types. These simply cannot be met under the current system.

In terms of the current AIT syllabus, the case for the computer-based examination as a replacement for the pen and paper response examination is diminished in that opportunity to express practical skills is of necessity limited and time constrained. Students in AIT typically work on rich tasks, often over several weeks, and the results of these may be extremely complex and employ skills and techniques which might only be briefly sampled during a three hour examination. However, almost all students indicated a preference for the assessment of practical performance at a computer with most indicating that the examination provided a sound and fair assessment of practical performance. They commented on the ease of working on the computer compared to working on paper citing correction of errors, speed of writing, amount of writing, speed of action and physical comfort whilst highlighting two major concerns: lack of time and worry about systems malfunctions. However, for the few students for whom this actually occurred almost immediate continuation was facilitated by moving to an alternative workstation.

The manner in which the examination was implemented was consistent, with the exception of the first instance, where the written section preceded the practical. Experience showed that the time devoted to this was considerably less than the one hour allocated and this led to invigilation problems. Typing into the *MS Word* document was efficient but one-hour was too long and the results in the end were of limited value. Students were fully occupied with the two hour practical section and putting this first made the examination more easily manageable. The performance tasks component of the exam provided scope for demonstration of capability. All students completed most requirements though some ran out of time and some lacked the skills required. Generally implementation was without technical difficulty apart from sound recording which, in spite of advanced warning, caused problems and disturbed the flow of the examination. Further, the sound recording was of limited value as most students just read what they had typed for the reflection.

In spite of the opportunity afforded to demonstrate capability, most students did not meet the high technical capability expected by their teachers. Many students didn't understand what a logo was and the inclusion of graphs was generally poor. Graphs were basic with little editing evident and no student created a standard temperature and rainfall composite chart which required two vertical axes. Less than 20 students designed a logo and used drawing tools to create it with most adapting one of the images supplied. However, overall the brochures were well done by many students with high visual appeal and some creative copy.

Student work resulting from the examination was able to be marked relatively reliably. The results of marking using the analytical rubric-based approach provided a good spread of scores with moderate correlations between the two markers (around $r=0.62$, $p<0.01$). There were also significant moderate correlations between the average assessors scores for the portfolio and for the examination ($r=0.58$, $p<0.01$). There was little correlation with any of these marks and those provided by the teachers except for the two schools that most rigorously implemented the portfolio. Most teachers did not provide a set of marks for the assessment task but rather a semester mark and grade. For the analytical marking of the performance tasks component of the examination Rasch analysis was completed that generated a reliable set of scores (Separation Index=0.85). This analysis found that for one of the criteria (creation of the logo) there was inconsistency in the use of the two highest levels of performance that was fixed by combining them. Notwithstanding the above, the examination was implemented consistently, and with minimal technical difficulty, was valid in terms of its match, though limited, to the AIT syllabus and typical pedagogy, and produced reliable scores when assessed by analytical marking.

The comparative-pairs approach to marking of 60 examples of the examination product, the brochure, provided a reliable set of scores (Separation Index=0.93) that was significantly correlated to the analytical marking scores of the examination both theory and practical sections ($r=0.73$, $p<0.01$). There were similar outcomes for rankings created by the two marking approaches.

As a replacement for the current pen and paper examination, the practical examination has strong credentials. Although limited in its scope, it does allow direct measurement of students' practical abilities and could be administered with the same technologies employed in the study with students' work samples being collected in on USB flash memory. However there are several areas of potential inconsistency between schools and these have implications for the standing of this type examination as set against examinations of other types and in other subject areas.

The first of these is the separation between screens of adjacent candidates. With current paper based examinations, a minimum separation is enforced by regulation. Even with similar separation, it is much easier to see a neighbour's work if it is upright and on a screen and some provision would have to be made for this difference even though in this study there was no evidence that this advantaged any student. Secondly, the communications potential of the standard workstation cannot be ignored. This might allow collusion between candidates, between candidates and external parties or information based web sites; a feature denied to all forms of paper based examination. In this study two school blocked access and in the others no student was observed to make use of the Internet. These issues will be revisited later in recommendations for future practice.

Digital Portfolio or Digital Examination?

In order to align with the aims, rationale, outcomes, content and preferred pedagogy, the AIT course assessment must include students using digital technologies. There are a number of ways in which this may be achieved and the research question effectively becomes, which method of assessment, portfolio or computer-based examination or combination, is most feasible for the course at this time? These two options were investigated because internationally they are the most likely to be used to assess students' capability in using ICT and teachers in Western Australia are familiar with both options. From the literature it was understood that each would have different strengths and weaknesses and thus a choice between them depends on the balance of positive factors and the options for addressing the weaknesses.

AIT is a relatively new course in Western Australia. The original assessment brief (Curriculum Council of WA, 2006) had provided for an assessment structure containing an electronic portfolio and a written examination of equal weighting. At that stage, the details of what would go into the portfolio and the exact format of the written examination had not been finalised and there was some disagreement over the examples that had been created. The main reasons why a portfolio had been recommended was the assumption that it would be easy to implement in the short term, would allow for the variety of contexts, would support the practical performance nature of the course and was already familiar to teachers in most of the subjects the course was replacing. It was not assumed that this would necessarily be the best long-term solution but that more time and research was required.

Then early in 2008 the decision was made that all students in Stage Two and Stage Three courses in Western Australia were to submit to an external examination with results being used to determine successful school graduation and ranking for entrance to tertiary institutions. It is to this decision that the current paper based examination owes its origin, for in the absence of any research data or tested alternatives, pen and paper became the fall back position, in spite of its limitations in courses such as AIT. The implications of the decision had widespread consequences including the problems associated with the suitability of an examination for lower achieving students. When an examination becomes a central focus of a course, the impact on course delivery is significant. Further, involvement of more students in the examination process gives rise to increased costs. Another confounding change was the requirement for the course to be packaged into a syllabus format with specific details of the content to be studied rather than what had been a definition of the boundaries of the content with the opportunity to address these to varying depths across contexts relevant to the students and teacher. The shift in focus towards content immediately highlighted the issue of the variety of relevant contexts that could be included in the course and the issue of the rapidly changing content in these areas of technology. With the focus on course outcomes, a range of contexts could be selected allowing adjustment to be made to content that might have become out-dated.

In general, from the data in the study, it may be concluded that *either* option, the digital portfolio or the computer-based examination, was able to be implemented successfully, although the examination was the easier to implement in a consistent fashion between schools. It appeared that only two of the seven teachers implemented the portfolio adequately according to the parameters agreed upon and this was reflected in the results, with students in these schools doing significantly better as a group. The only implementation issues for the examination were the failure of audio recording in three schools and the handful of students who experienced minor technical disruption at a cost of no more than a few minutes.

While the examination was well implemented from a technical perspective, analytical marking was not as reliable as for the portfolio although correlation between markers was good and statistically significant for both methods. The comparative-pairs marking process, used for the examination practical, was also highly reliable. The main reason for lower reliability for the analytical marking of the examination appeared to be because on a few criteria, very few top marks were given, for example for correct file formats, logo and brochure, explanation or justification. Rasch analysis of the portfolio analytical marks (Product, Process Document and Extra Artefacts analysed separately) gave a reliable set of scores for all three components. No modifications were required although for the Extra Artefacts the thresholds for three of the criteria did not work very well. The analytical marks generated by the three components of the Portfolio were not highly correlated probably indicating that they were addressing different types of performance.

A major weakness of the examination was that a relatively low level set of tasks were required (logo, brochure, and spreadsheet graph) in order that students in all classes could attempt them using a typically standard set of software that they would all have access to. The types of practical tasks which teachers give students vary greatly, for example many teachers do not include database related tasks or even spreadsheets while others will not do animations or movie editing. Therefore, it is very difficult to set examination tasks that all students would have the background to attempt and that would also allow adequate scope for the more capable students to demonstrate their ability. This is not a problem with the portfolio, indeed it is an advantage, as the design brief could be varied between classes and typically allow tasks to be relatively open-ended. However, consistent assessment of such disparate submissions may later present difficulties.

In summary, the portfolio may be easily implemented by teachers whilst requiring strict invigilation procedures to be adhered to. Students’ work may be burned to disk or copied to flash memory for submission. No major technical issues present, though some difficulty might occur in marking the diverse range and size of files submitted. Teachers may set tasks appropriate to the technology available. Inconsistency of implementation could be a problem. Portfolios allow good discrimination to be made between students and are familiar and well understood.

Computer-based examinations are easy to implement consistently by either teachers or external invigilators. Currently they may be simply and reliably delivered and collected using USB flash memory. On site testing would be required in advance to ensure all required technologies work. It is difficult to ensure that students cannot view each other’s screens and difficult to set tasks that all students can attempt and that, at the same time, would extend more capable students. Quite good discrimination between students is achievable given appropriate assessment tasks.

Recommendations

This section makes recommendations, based on an analysis of the data, for the universal implementation of an electronic portfolio and/or a computer-based examination in the Stage Two AIT course in Western Australia.

While students may include the successful study of AIT towards secondary high school graduation or tertiary entrance, this would be of little value if the form of external assessment propels the course towards becoming mainly 'book work' rather than creative digital work. We are living in a society where almost every avenue of work and life requires the use of digital tools and resources. Whether a senior student is aiming to be a mechanic, doctor, accountant or travel agent, study in AIT could begin to give them the skills, attitudes and understanding that would support them in being more successful in work and life. There are a number of ways in which students could be assessed on their use of digital technologies and this study has looked at two: the digital portfolio and the computer-based examination.

Overall when comparing the Portfolio and the Examination there was no compelling reason to choose one over the other, each had strengths and weaknesses. Therefore it is recommended that any decision be made on the basis of how well each could be implemented in the manner outlined in the following section. The choice is a digital portfolio, a computer-based examination or perhaps a combination of each.

Implementing a Digital Portfolio

If the Portfolio were to be implemented the structure used in the study is recommended. This allows students to adequately demonstrate their capability with some scope for tailoring to the context for the student. However, ideally it would need an online portfolio management system and would need a well-structured system for verification that would probably include some type of signed affidavit with spot checks on a sample of students to ensure all teachers implemented the portfolio according to the required procedures and conditions.

The portfolio could consist of three parts: a digital product created over about 15 hours, a process document collating research and development of the product created over about 5 hours and the submission of two previously created digital artefacts, with short descriptions of the development process, over about one hour. A set of clearly written parameters needs to be set for the Portfolio starting with the time limits for development of each section as described above.

A choice of design briefs would also be needed. Although ideally teachers would be allowed to set their own many teachers might not feel confident enough to do this and thus three or four example design briefs should be supplied for them to use or modify with the proviso that conformity to the appropriate design brief parameters is maintained. These would specify the aim, purpose, included components, file sizes and file types allowed.

The process document should consist of four sections: Research, Design, Production and Evaluation with suggested limits for each and consist of the selection and collation of material resulting from the development of the digital product best representing the process of development employed.

The nature of the two additional artefacts should allow for demonstration of skills in areas other than those demonstrated in the main digital product. Each artefact should be supported by a brief description of the hardware and software employed, the design principles and conventions displayed, the skills demonstrated and make reference to any assistance received.

The components of the Portfolio need to be marked separately using different criteria as they represent different types of performance. For example, the Process Document provides an opportunity to demonstrate an understanding of the Technology Process and capability in planning, analysing, organising, managing and evaluating.

Implementing a Computer-Based Performance Examination

If the performance tasks examination were to be implemented then the structure used in the study is recommended with some minor modifications, principally the removal of audio reflection. The study has highlighted two areas in which decisions would need to be made: technical implementation; and performance tasks specification.

Technical Implementation

The project used USB flash memory that worked on all school computers for all students. This would be cumbersome but not unrealistic to scale up to state-wide implementation but in the long-term an online examination management system should be used. However, this introduces management of the internet service by schools as a variable. Schools in the study had different levels of network management and support and the study found that consistency of service and support could not always be relied upon.

A further related question is whether access to the Internet should be precluded? In this study it was not a requirement although at least two schools did this of their own accord. There was a requirement that students could only use the 18 digital photographs that were provided on the USB flash drive so there was no incentive to search for additional media.

Another question is whether software should be specified. In this study, no attempt was made to limit access to software that was normally available to the students. The basic nature of the tasks meant that this provided little, if any, advantage to any students with most using office application software and a relatively basic graphics package. The choice would seem to be to specify the software allowed or to allow all available or to provide software on the USB flash memory in a system similar to that employed by Fluck et al. (2009). In the latter a highly consistent and secure environment was created by the temporary installation of both the operating system and software on students' own laptop computers. The option of paper-based or digital design phase appeared to be appreciated by students with most opting for paper-based design and this should be allowed for even though scanning of these introduces an additional administrative burden.

Performance Tasks Specification

In the study the performance tasks were selected to be relatively low-level and easy to ensure some degree of engagement for all students. Further these tasks were defined fairly explicitly with little or no choice for students. Clearly if this were done for the final external assessment then it would limit the opportunity for high ability students and would tend to stifle context variety in the course. Therefore it is likely that some degree of choice of context, type of solution and/or tools used is needed. However, this makes the development of appropriate tasks, description of those tasks and the statement of assessment criteria much more difficult to generate. Given that currently many teachers do not appear to be addressing the content as specified in the syllabus, with some focusing almost exclusively on interactive multimedia development and others on business software, the examination would have to follow a pre-defined structure so that teachers could adjust their teaching programmes accordingly. Advance warning of the skills required in the examination would need to be circulated to schools informing students that, for example, they will have to develop a graphic, edit a photograph, create a spreadsheet, access a database and hyperlink media elements. What would then be unknown would be the specific scenario to be addressed and the associated contextual media provided. All of these factors would need to be further investigated with a more complex task, allowing for more choice.

Future Directions for Digital Forms of Assessment

Digital technologies have already transformed the world for which we are educating today's children. It is a world of social networks, on demand services and real time event reporting. Today's students live in a different world from that of their predecessors. In their working lives they will not be required to memorise or carry enormous amounts of information around with them; they will need only to quickly access and organise information from multiple sources and to be able to check the reliability of this information. It follows that assessment of skills such as factual recall, beyond the basics of literacy and numeracy, are already largely redundant and must inevitably be succeeded by assessment of the higher order thinking skills of analysis, synthesis and evaluation. Current examinations mainly test knowledge and some understanding, typically in a room devoid of technology using a pen, a sheet of paper and a three hour time limit. The overemphasis on low-level thinking, to the detriment of imagination and creativity, is plainly inauthentic and anti-educational. What then might assessments of the future look like?

In the UK, the end of pen and paper testing has been predicted by Isabel Nisbet, chief executive of *Ofqual* the regulator of qualifications, examinations and assessments in England and Wales She was quoted as saying that there was now an,

> issue of validity with traditional paper exams because pupils no longer wrote things to assist with learning. We need to make sure [the way pupils are tested] isn't overtaken by the modern world and doesn't become a relic of the early 20th century (Shepherd, 2010).

However Dylan William, professor of educational assessment at the Institute of Education, University of London, is quoted as predicting that it would be 20 years before all public examinations were taken at computers.

> The skills needed to sit an exam at a computer are different from those required for a pen-and-paper test... it would be logistically difficult for schools to administer computer-based tests for lack of technology.

> Subjects such as Mathematics might be particularly problematic because "current software does not make it easy for students to draw graphs or write formulae. I think we are probably about 20 years away from having all exams sat at computer terminals" (Shepherd, 2010).

In Denmark, a leading advocate and early adopter of digital technologies in assessment, the Ministry of Education has moved from allowing computers to be used to type examination responses to a 2 ½ year pilot project trialling ICT based examinations in six subjects with access to internet allowed. A total of 14 colleges are piloting the new system and all schools in the country have been invited to join the scheme by 2011. Hobson (2009) describes a typical scenario.

> On the morning of the exam, the exam room the floor is covered in cables. IT experts are busy helping the teenagers set up their laptops, making sure they all work. At five to nine, the room falls silent. CD-ROMs and exam papers are handed out together. This is the Danish language exam. One of the teachers stands in front of the class and explains the rules. She tells the candidates they can use the internet to answer any of the four questions. They can access any site they like, even Facebook, but they cannot message each other or email anyone outside the classroom. At nine o'clock the exam begins.

Hobson (2009) quotes Sanne Yde Schmidt the head of the project at Greve High School on the question of potential cheating. "The main precaution is that we trust them. I think the cheat rate is very low because the consequences of cheating are very big." Students admit that cheating is possible but that they are unlikely to do it. Hobson (2009) quotes Pernille Günther Jensby and Nina Ahmed on this topic. “It's possible to cheat but I think we have so much respect and self-discipline, so we won't do it." "I think it'd be very difficult [to cheat] because you don't have time, you're under pressure, and you have too many tasks."

The type of questions asked also makes cheating superfluous with those requiring recall of facts no longer included on the paper. The new questions call for higher order skills with the emphasis on quick selection and analysis information. The following quotation by Bertel Haarder, Danish Minister for Education, is from Hobson (2009).

> Our exams have to reflect daily life in the classroom and daily life in the classroom has to reflect life in society. The internet is indispensible, including in the exam situation. I'm sure that is would be a matter of very few years when most European countries will be on the same line.

Conclusion

This study has added new knowledge to the fields of Assessment of Human Performance and Computer Assisted (or Digital) Assessment. Specifically, in the case of AIT in Western Australian schools, two forms of performance based assessment have been shown to be viable, potential replacement candidates for existing paper based examinations. Though each has its own limitations, improvements in assessment validity and authenticity are without question, and reliability of scoring, though acceptable, could easily be improved upon for both a digital portfolio and a computer based examination.

The study has also added new knowledge in the area of marking methods. Comparative-pairs marking of digitally created and web-accessed students' work samples has been shown to be highly reliable and no-more difficult or time consuming than rubric based scoring. It appears to be a superior method where the assessment task is holistic and mature.

Areas of further research highlighted by this study are many fold. The potential for technological failure and the anxiety for students caused by hardware/software failure, though of minimal impact in this study, remains perhaps the greatest obstacle to adoption of digital assessment. The absence of a standard operating environment and differences in computer hardware and application software may also be seen as obstacles to the fairness and equity of digital assessment. Both of these factors point to web-based solutions and it is here that further research seems likely to be directed. A web-delivered digital assessment could allow a standardisation of the assessment environment and application software; with students working on 'cloud based' applications, the local machine would serve merely as a connection interface or terminal.

Rather than merely refining current methods and well-established techniques or questioning the reliability or validity of a specific approach, what is surely required is a re-definition of the *suitability* of the assessment for the *purpose* in question. This calls into question many of the assumptions and beliefs on which the current examination system is predicated.

> Belief in the power of conventional summative assessment techniques to be objective and efficient, to motivate present performance and to predict future performance is being challenged by a range of research evidence that identifies the significant flaws in these assumptions. Moreover, the assumptions highlight the worrying price that the use of assessment to measure and control extracts, including reduced motivation and significantly lower performance on the part of students (Broadfoot & Black, 2004, p.8).

REFERENCES

Ahmed, A. & Pollitt, A. (2001). Improving the validity of contextualised questions. University of Cambridge Local Examinations Syndicate. *Paper presented at the BERA Conference, Leeds, September 2001*. Retrieved June 2011 from http://www.cambridgeassessment.org.uk/ca/digitalAssets/113953_Improving_the_Validity_of_Contextualised_Questions.pdf

Allen Consulting Group (2006). *World class skills for world class industries*. Report to the Australian Industry Group. Retrieved June 2011 from http://www.aigroup.com.au/portal/binary/com.epicentric.contentmanagement.servlet.ContentDeliveryServlet/LIVE_CONTENT/Publications/Reports/2006/World_Skills_Full.pdf

CISCO (2010). *Assessment & teaching of 21st century skills.* Retrieved June 2011 from http://www.cisco.com/web/about/citizenship/socio-economic/docs/ATC21S_Exec_Summary.pdf

Barrett, H. C. (2007). Researching electronic portfolios and learner engagement: the reflect initiative. *Journal of Adolescent & Adult Literacy*, *50*(6), 436-449.

BECTA (2006). *E-assessment and e-portfolios.* UK, Coventry: Becta's View. Retrieved June 2011 from http://www.excellencegateway.org.uk/page.aspx?o=138059

Beevers, C. E., Foster, M. G., McGuire, G. R. & Renshaw, J. H. (1992). Some problems of mathematical CAL. *Computers & Education*, *18*(1-3), 119-125.

Beetham, Helen (2005). *E-portfolios in post-16 learning in the UK: developments, issues and opportunities.* Report prepared for the JISC e-Learning and Pedagogy strand of the JISC e-Learning Programme. Retrieved June 2011 from http://www.jisc.ac.uk/media/documents/themes/elearning/eportfolioped.pdf

Biggs, J.B. & Moore, P.J. (1993). *The process of learning.* Sydney: Prentice Hall .

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (2000). How people learn - Brain, mind, experience and school. Washington, D. C.: National Academy Press.

Broadfoot, P. & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education: Principles, Policy & Practice, 11*(1), 7-26.

Brown, S. and Knight, P. (1994). *Assessing learners in higher education.* UK, Oxon: Routledge Falmer (ISBN 0 7494 1113 9).

Brown, G., Bull, J. & Pendlebury, M. (1997). *Assessing student learning in higher education.* UK, London: Routledge (ISBN 0 415 16226 2).

Brown, G. (1968). Role of examinations in society. *Conference proceedings on examinations. Queens University of Belfast.*

Bull, Joanna and McKenna, Colleen (2004). *A Blueprint for computer-assisted assessment.* UK: Routledge (ISBN: 978-0-415-28704-3).

Burns, R. B. (1996). *Introduction to research methods.* Australia, South Melbourne: Addison Wesley Longman.

Campbell, A. (2008). *Performance enhancement of the task assessment process through the application of an electronic performance support system*. Theses: Doctorates and Masters, Edith Cowan University, Perth, Western Australia. Retrieved June 2011 from http://ro.ecu.edu.au/theses/18

Carney, J. (2004). Setting an agenda for electronic portfolio research: A framework for evaluating portfolio literature. *Presentation at the American Educational Research Association Conference, April 14, 2004.* Retrieved June 2011 from http://it.wce.wwu.edu/carney/Presentations/AERA04/AERAresearchlit.pdf

Cole, M. (1990). Cognitive development and formal schooling: The evidence from cross-cultural research. In L. C. Moll (Ed.), *Vygotsky and education: Instructional* *implications of sociohistorical psychology*, 89-110. New York: Cambridge University Press.

Colley, K. (2008). Performance-based assessment. *The Science Teacher, 75*(8), 68-72. Retrieved October, 2010, from Academic Research Library (ID: 1596595471).

Crawford, K. & Fakete, A. (1997). *What do exams really measure?* Proceedings of the 2nd Australasian conference on Computer Science Education, 185-190 (ISBN:0-89791-958-0).

Cronbach, L. J. (1990). *Essentials of psychological testing.* New York: Harper Collins.

Curriculum Council of Western Australia (2009). *Applied Information Technology course syllabus.* Retrieved June 2011 from http://www.curriculum.wa.edu.au/internet/Senior_Secondary/Courses/WACE_Courses/Applied_Info_Tech

Curriculum Framework of Western Australia (1998). *Principles of Learning Teaching and Assessment.* Retrieved June 2011 from http://www.det.wa.edu.au/professionallearning/detcms/professionallearning/professional-learning-website/for-teaching-staff/graduate-teacher/principles-of-learning-teaching-and-assessment.en?page=2&tab=Main#toc2

Darling-Hammond, L. & Anderson, F. (2010). *Beyond basic skills: The role of performance assessment in achieving 21st century standards of learning.* Stanford Centre for Opportunity Policy in Education. Retrieved June 2011 from http://edpolicy.stanford.edu/pages/pubs/pub_docs/assessment/scope_pa_overview.pdf

Elliott G. & Greatorex, J. (2002). A fair comparison? The evolution of methods of comparability in national assessment. *Educational Studies, 28*(3), 253-265.

Educational Testing Services (2002). International ICT Literacy Panel. *Digital transformation: A framework for ICT literacy.* Princeton, N.J.: Educational Testing

Service. Retrieved June 2011 from
http://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf

Fisette, J. L., Placek, J. H., Avery, M., Dyson, B., Fox, C., Franck, M., Graber, K., Rink, J. & Zhu, W. (2009) Developing quality physical education through student assessments. *Strategies: A Journal for Physical and Sport Educators, 22*(3) 33-34.

Fitzpatrick, R. & Morrison, E.J. (1971). Performance and product evaluation. In E.L. Thorndike (Ed.) *Educational Measurement (2nd ed.),* 237-270. Washington, D.C.: American Council on Education.

Fluck, A., Pullen, D. & Harper, C. (2009). Case study of a computer-based examination system. *Australian Journal of Educational Technology, 25*(4), 509-523.

Hattie J (2009) *Visible Learning; a synthesis of over 800 meta-analyses relating to achievement.* London; Routledge

Hammann, M., Hoi Phan, T. T., Ehmer, M. & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education, 42*(2), 66-72.

Hobson, J. (2009, November 4). Danish pupils use web in exams. Retrieved June 2011 from http://news.bbc.co.uk/2/hi/uk_news/education/8341886.stm

Husen, T. (1994). Research paradigms in education. In Husen, T. & Postlethwaite, T.N. (Eds.) *The International Encyclopaedia of Education, 9,* 5055. New York: Elsevier Science Ltd.

International Test Commission (2005). *International guidelines on computer-based and internet delivered testing.* Retrieved June 2011 from
http://www.intestcom.org/_guidelines/guidelines/index.html

JISC (2007). *Effective practice with e-Assessment.* Retrieved June 2011 from
http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf

Kennedy, D.M. (2008). Digital literacy research. In L.A. Tomei (Ed.) *Information Technology Curriculum Integration, 1,* 228. New York: Hershley.

Kimbell, R.A., Wheeler, A., Miller, S. & Pollitt, A. (2007). *E-scape portfolio assessment (e-solutions for creative assessment in portfolio environments) phase 2 report.* TERU, Goldsmiths University of London (ISBN 978-1-904158-79-0).

Kinshuck, Chen, N. S. & Lin, K.M. (2008). Analysing users' satisfaction with e-learning using a negative critical incidents approach*. Innovations in Education and Teaching International, 45*(2), 115-126.

Knight, P. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education, 27*(2), 275-286.

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

Lane, S. (2004). Validity of high-stakes assessment: Are students engaged in complex thinking? *Educational Measurement: Issues and Practice*, *23*, 6-14.

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

Linn, R.L. and Gronlund, N.E. (1995). *Measurement and assessment in teaching (7th ed.).* NJ Englewood Cliffs: Prentice-Hall.

Linn, R.L., Baker, E.L. & Dunbar S.B (1990). *Complex, performance-based assessment: Expectations and validation criteria.* CSE Technical Report 331.Retrieved June 2011 from http://www.cse.ucla.edu/products/Reports/TECH331.pdf

Marton, F. & Säljö, R. (1976). On qualitative differences in learning: Outcome and process. *British Journal of Educational Psychology, 46*, 4-11.

McGaw, B. (2006). Assessment fit for purpose. *Paper presented at the 32nd Annual Conference of the International Association for Educational assessment, Singapore.*

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

Messick, S. (1996). Validity of performance assessment. In G. Philips (Ed.), *Technical Issues in Long-Scale Assessment.* Washington, D.C.: National Centre for Educational Statistics.

Newble, D.I. and Jaeger K. (1983*).* The effect of assessments and examinations on the learning of medical students. *Medical Education, 17,* 165-171.

MCEETYA. (2005). *National assessment program ICT literacy years 6 &10 report 2005.* Carlton South, Australia: Curriculum Corporation. Retrieved June 2011 from http://www.nap.edu.au/_Documents/MCEECDYA/2005%20ICTL%20Public%20Report.pdf

Parshall, C.G., Spray, J.A., Kalohn, J.C. & Davey, T. (2002). *Practical considerations in computer-based testing.* New York: Springer-Verlag (ISBN0-387-98731-2).

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment Research & Evaluation, 13*(4), 1-11. Retrieved June 2011 from http://pareonline.net/getvn.asp?v=13&n=4

Pellegrino, J. W., Chudowsky, N. & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Pollitt, A. (2004). Let's stop marking exams. *Paper presented at The International Association for educational Assessment Conference, Philadelphia.* Retrieved June 2011 from
http://www.cambridgeassessment.org.uk/ca/digitalAssets/113942_Let_s_Stop_Marking

_Exams.pdf

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon.* NCB University Press, *9*(5), Retrieved June 2011 from http://www.marcprensky.com/writing/prensky%20-%20digital%20natives,%20digital%20immigrants%20-%20part1.pdf

Qualification Curriculum Authority (2007). Regulatory principles for e-assessment. Retrieved January 2012 from www.sqa.org.uk/files_ccc/RegulatoryPrinciplesforE-assessment.pdf

Qualification Curriculum Authority (2008). *Key stage 3 ICT assessment tasks.* Retrieved June 2011 from http://webarchive.nationalarchives.gov.uk/20090810204322/testsandexams.qcda.gov.uk/18908.aspx

Resnick, L. B. & Resnick D. P. (1992). *Assessing the thinking curriculum: New tools for educational reform.* In B. R. Gifford and M. C. O'Connor (Eds.) *Changing assessments: Alternative views of aptitude, achievement and instruction,* 37-75. Boston: Kluwer Academic Publishers.

Ridgway, J., McCusker, S. and Pead, D. (2004). *Literature Review of E-assessment.* Report10 Futurelab Series. Retrieved June 2011 from http://www.futurelab.org.uk/research/reviews/10_01.htm

Romeo, G. (2008). *Information and communication technologies in education*. In Yelland, N. Greg, A. & Dakich, E. (Eds.) *Rethinking Education with ICT. 14*, 214-215. Rotterdam: Sense Publishers,

RUMM Laboratory Pty Ltd (2011) http://www.rummlab.com.au/

Salvia, J. & Ysseldyke, J. E. (1998). *Assessment (7th Ed.)*. Boston: Houghton Mifflin.

Schmidt, H. G., Norman, G. R. & Boshuizen, H. P.A. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine*, *65*, 611-621.

Scriven, M. (1967). The methodology of evaluation. In Tyler, R. W., Gagne, R.M. & Scriven, M. (Eds.) *Perspectives of curriculum evaluation*. Chicago: Rand McNally.

Shepherd, J. (2010, August 18) Exams: changing habits may spell end for pen-and-paper tests. Retrieved June 2011 from http://www.guardian.co.uk/education/2010/aug/18/exams-keyboard-answers-ofqual

Spector, J. M. (2006). A methodology for assessing learning in complex and ill-structured task domains. *Innovations in Education and Teaching International, 43*(2), 109-120.

Taylor, A. R. (2005). *A future in the process of arrival. Guidelines for the development and use of computer-based assessments.* UK, Leicester, P. T. Centre: The British Psychological Society.

The British Psychological Society (2002). *Guidelines for the development and use of computer-based assessments.* Retrieved June 2011 from http://www.psychtesting.org.uk/downloadfile.cfm?file_uuid=64877B7B-CF1C-D577-971D-425278FA08CC&ext=pdf

The Council of the International Test Commission (2005). *International guidelines on computer-based and internet delivered testing.* Spain, Granada: The Council of the International Test Commission. Retrieved June 2011 from http://www.intestcom.org/Downloads/ITC%20Guidelines%20on%20Computer%20-%20version%202005%20approved.pdf

Thomas, P,R. & Bain, J.D. (1984). Contextual dependence of learning approaches: The effects of assessments. *Human Learning*, *3*, 227-240.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34,* 273–286.

Tomlinson, M. (2004). *14-19 Curriculum and qualifications reform: Interim report of the working group on 14-19 reform*. London: DfES.

Vernon, D.T. & Blake, R.L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine, 68*, 550-563.

Wadsworth, Y. (1998). What is participatory action research? *Action Research International, Paper 2.* Retrieved June 2011 from http://www.scu.edu.au/schools/gcm/ar/ari/p-ywadsworth98.html

Watkins, D. & Hattie, J. (1985). A longitudinal study of the approaches to learning of Australian tertiary students. *Human Learning, 4,* 127-141.

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment Research and Evaluation, 2*(2), Retrieved June 2011 from http://pareonline.net/getvn.asp?v=2&n=2

Willig, C. (2001). *Introducing qualitative research in psychology adventures in theory and method.* UK, Buckingham: Open University Press.

Zane, T.W. (2009). Performance assessment design principles gleaned from constructivist learning theory (part 2). *TechTrends*, *53*(3), 86-94. Boston: Springer.

APPENDIXES

Appendix A

AIT Course Rationale and Course Outline

Information and communication technologies (ICT) are rapidly changing the way we live and work in Australia. They affect the nature of communication, entertainment and lifestyle decisions. Almost every area of employment requires some understanding and use of ICT for greater productivity and creativity. Every school graduate needs to be capable of using ICT in his/her personal, community and future professional lives. This course aims to address that need, by providing students with creative opportunities through interesting practical experiences, using exciting and innovative software and equipment.

These technologies are increasingly becoming part of everything we do within a knowledge-based society, built around the innovative, creative and enterprising use of ICT to improve the standard of living. All Australians need to possess and be empowered by understanding, experience and skills in the nature and use of ICT. This course aims to equip post-compulsory secondary students for current and future study (TAFE or University), employability and successful inclusion in a modern society.

The course focuses on the application of computer technologies to living in the community and working in industry and business environments. It looks at the impact on workplaces, individuals and society. As such, it provides opportunities for students to develop knowledge and skills relevant to the use of ICT to meet everyday challenges. Students consider a variety of computer applications for use in their own lives, business and the wider community. They consider the ethical implications of ICT solutions and develop an appreciation of the role and impact of these technologies on their personal values, and those within a democratic and ethnically diverse society.

An integral aspect of the course is the opportunity to address vocational competencies, leading to the possibility of a range of VET accreditation, thereby contributing to the international employability of students.

The emphasis of the course is on developing an understanding, from a user's perspective, of the application of computer technology in various contexts and the design of information solutions to meet challenges encountered in those contexts. This will require an understanding of the nature of these challenges and contexts as well as associated work processes.

Students build their understanding, experience and skills by investigating, designing, constructing and evaluating ICT solutions, using a variety of software applications, including some commercial applications commonly used in business and home environments. They consider such solutions within personal, community and workplace environments. Students gain essential life and work skills in problem-solving, time management and communications skills, while working both independently and collaboratively. The course provides an excellent general grounding in ICT for the future study aspirations and professional lives of all students.

This course provides students with the opportunity to further their achievement of specific overarching learning outcomes from the Curriculum Framework together with the development of the core-shared values.

The course is designed to facilitate the achievement of four outcomes. These outcomes are based on the Technology and Enterprise learning area outcomes in the Curriculum Framework. Outcomes are statements of what students should know, understand, value and be able to do as a result of the syllabus content taught.

Outcome 1: Technology process

Students apply a technology process when creating or modifying information solutions using information communication technologies. In achieving this outcome, students:

- investigate ideas considering alternatives;
- devise, communicate and evaluate proposals and design plans in appropriate forms; and
- implement and evaluate production processes and strategies to manage resources efficiently.

Outcome 2: Understanding information and communication technologies

Students understand the nature and use of computer hardware and software to achieve information solutions. In achieving this outcome, students:

- understand the ICT-related concepts, formats and terminology required to select and use
- appropriate software and hardware to achieve information solutions;
- understand the relationship between forms, structures and conventions of information solutions that influence the selection and use of ICT; and
- understand management, processes, procedures and techniques required to achieve information solutions.

Outcome 3: Quality of information solutions

Students explore alternatives and use skills, techniques, processes, standards and conventions to achieve information solutions.

In achieving this outcome, students:

- apply appropriate forms, structures and conventions to create or modify information solutions;
- apply skills, techniques and processes to develop information solutions; and apply enterprising capabilities, exploring alternatives, in working to achieve information solutions.

Outcome 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.

In achieving this outcome, students:

- understand the cultural beliefs, values, abilities and ethical positions that can impact on the use of ICT
- understand the consequences of ICT use in different contexts and how this relates to beliefs, values, abilities and ethical positions; and □
- understand the consequences of technological developments on structures and environments.

UNIT 2BAIT

Unit description

The focus for this unit is **information and communication technologies in business**. Skills, principles and practices associated with various types of businesses to enhance students' career prospects are emphasised. Students examine the use of ICT in a range of administrative and business environments. They identify and explain the components and configuration of a computer system to meet the needs of the organisation. Students design information solutions for problems encountered in these contexts and understand the social issues inherent in work practices.

Unit learning contexts

Within the focus area of **information and communication technologies in business**, teachers may choose one or more of the following contexts (this list is not exhaustive):

- small commercial business
- large commercial business
- government organisation
- non-profit organisation.

Unit content

This unit includes knowledge, understanding and skills to the degree of complexity described below:

Social implications and trends

The impact of ICT on individuals, communities, and environments

- describe the impacts of business technologies on work expectations and lifestyles:
- mobile technology devices on business e.g. notebooks, mobile and smart phones, PDA's
- 24/7 communication expectations
- analyse the impacts and effects of local and national ICT structures on the flow of information to specific communities and environments:
- mobile phone networks
- internet/broadband
- wireless technologies
- Internet cafes and libraries.

Values, ethics and inclusivity

- examine ethical issues related to the use of technologies within business:
- identity theft e.g. phishing, pharming
- piracy i.e. software
- breach of intellectual property i.e. unauthorised use of programming code
- breach of copyright e.g. illegal file sharing, bootleg recordings, screener DVD's, plagiarism
- security issues i.e. hackers
- discuss the impacts of the global digital divide on business e.g. e-commerce—paying bills online, online shopping.

Past and emerging trends in ICT

- examine the emergence of electronic commerce and global networks:
- banking—past (counter banking) to current (online banking),
- paying bills—past (paying by cheque) to current (BPay)
- business accounts—past (account ledgers and books) to current (accounting software)
- investigate effects of business globalisation design and manufacture of information products in different countries:
- outsourcing of ICT products—define, examples, advantages/disadvantages.

Hardware and software

Hardware components and functions

- define networking concepts:
- physical transmission media (network cable and wireless connection)
- basic network components e.g. server, router, modem, network interface card, switch
- security of networks i.e. firewalls, passwords
- describe connection issues relating to the following components:
- input/output devices e.g. Bluetooth, USB, wireless, infrared
- communication devices e.g. PDA, mobile phone.

Applications and systems software

- describe the application/use of common ICT business software including descriptions, examples and use of:
- personal information managers i.e. Outlook, Lotus Notes
- presentation software for business
- word processing
- simple spreadsheets—basic formulas and charting
- flat file databases
- business Office Suites i.e. MS Office, Open Office
- online office applications i.e. Google applications
- publishing e.g. Adobe suite
- online forms for data collection.

Design and acquisition of hardware and software

- introduction to the following factors when purchasing/acquiring hardware/software in a business environment:
- cost versus benefits in a business environment
- ergonomic requirements e.g. workstations. Does the equipment meet Occupational Health and Safety requirements?
- legal acquisition of hardware/software i.e. license types: freeware, shareware, proprietary—have you purchased a legal copy/licence of the software?
- warranty and technical support.

Digital data and information

The nature, forms and transfer of digital data

- investigate various solutions for business documents used in print and online business environments:
- transfer of data files between ICT environments using email, CD/DVD ROMS, flash drives
- file compression, conversion and size e.g. lossy, lossless, zip
- security of data e.g. pdf and flash documents
- introduction to encryption including the use of public and private keys.

Processing and managing data

- incorporate input validation rules for text, numerical and image based data into business products
- manipulation of data e.g. saving data in various formats
- introduction to knowledge management systems and their use in business e.g. capture sales data and produce reports
- input and extract data from a data store (database, Information Management system, email system).

Creative application of information design principles

- define and apply information design principles in the creation of business related technology products considering:
- a specified target audience including:
- intent/purpose
- content
- presentation medium
- aesthetics
- relevant language and terminology
- typography
- creative design principles including:
- balance—symmetry, asymmetry, radial
- proportion
- dominance
- harmony/unity/proximity
- elements of design including:
- space—positive and negative
- texture
- colour
- line
- alignment/position
- shape/form.

Workplace, practices and careers

Careers, work and jobs

- demonstrate an awareness of training opportunities in relevant business software
- investigation of alternative career pathways for business employment e.g. data entry and operator, digital accounts
- show an understanding of multi-skilling.
-
- Work environments and legislation
- investigate the impact of ICT use within a business organisation:
- ICT code of conduct i.e. appropriate use of equipment and ethical use of company information, email and internet usage, proxy server logs
- introduction to company liability in terms of:
- Occupational Safety and Health
- employee privacy.

Technology processes in the workplace

- analysis of client requirements and needs i.e. requirements analysis
- describe factors influencing teams in a business environment i.e. job security, gender bias, communication.

Assessment

The four types of assessment in the table below are consistent with the teaching and learning strategies considered to be the most supportive of student achievement of the outcomes in the Applied Information Technology course. The table provides details of the assessment type, examples of different ways that these assessment types can be applied and the weighting range for each assessment type.

Weighting Stage 2	**Type of assessment**
10–20%	**Investigation** Research work in which students plan, conduct and communicate an investigation. Investigation of ICT-related issues or cultural contexts, exploring a range of primary and secondary sources. *Best suited to the collection of evidence of student achievement of Outcomes 1, 2 and 4.*
40–50%	**Production /performance** Extended production project in which students explore ideas and control the processes required to manage the quality of production. Students engage in an activity or on-the-spot evaluation of a performance. This may be one large production /performance task or it may be two or more smaller tasks. Manage a range of production processes, evaluating and modifying them as necessary. Demonstrate an understanding of styles, structures, codes and conventions and the development of confidence and competence in the use of technologies, skills and processes in a range of contexts. *Best suited to the collection of evidence of student achievement of Outcomes 1 and 3 and includes using a journal to show evidence of exploration and the development of ideas, reflection on learning processes and critical evaluation and modification of ideas.*
10–20%	**Response** Students apply their knowledge and skills in analysing and responding to a series of stimuli or prompts. Response to, analysis and evaluation of own or professional information technology products. Types of evidence may include: observation checklists, journal and evaluation tools (self or peer). *Best suited to the collection of evidence of student achievement of Outcomes 2 and 4.*
10–20%	**Examination** Students apply their understanding and skills in Applied Information Technology to analyse, interpret, solve problems and answer questions in examination settings. Students are required to use technical terminology, apply knowledge and application of Applied Information Technology skills. Types of questions will include multiple choice, short answer, extended answer and production. *Best suited to the collection of evidence of student achievement of Outcomes 1, 2, 3 and 4.*

Appendix B

Situation Analysis

It was important that assessment tasks constituted good professional practice, met the requirements of the course and were reasonably able to be implemented by a 'good' teacher in a real school. A situation analysis was required to consider what was possible within the requirements of the course and the constraints of the school environment and teacher and student capabilities. However, the aim was to move towards the 'cutting edge' of what was possible.

The Students (1) What are key characteristics of typical students? (2) How do these characteristics relate to assessment in the course? (3) How do these characteristics relate to the use of ICT?	Year 11 doing Unit 2B (Business context) likely to be reasonably ICT capable Students from 'real' schools? Country school? Range of schools with multiple classes. Students have expectations of working in an IT environment so assessment should match. Does this match the expectation of teachers? Some students do the course to gain functional literacy, some have interest in IT, some want a career involving ICT, some are aiming for TAFE or Uni (portfolio entries), some are connected to VET competencies.
The Course (1) What are the main intended outcomes? (2) What are the main pedagogies? (3) What are the assessment requirements in terms of knowledge and skills or capability? (4) How is external assessment conducted? (5) How is school-based assessment conducted?	The syllabus is currently being refined – mainly affect content. Specific contexts are referred to in the syllabus. The course concerns IT skills, content knowledge, product development, values (related to main outcomes). Assessment types: Production/Performance (50-60%), Investigation (20-30%), and Response (20-30%). Currently assessed externally through m/c, short answer, extended answer and written design problem. School based assessment is mainly conducted through applied tasks using a variety of ICT. Also through response activities (e.g. tests, assignments). Content is connected to the contexts the teacher selects for the tasks. **Unit 2BAIT** The focus for this unit is information and communication technologies in business. Skills, principles and practices associated with various types of businesses to enhance students' career prospects are emphasised. Students examine the use of ICT in a range of administrative and business environments. They identify and explain the components and configuration of a computer system to meet the needs of the organisation. Students design information solutions for problems encountered in these contexts and understand the social issues inherent in work practices. Within the focus area of information communication technologies in business, teachers may choose one or more of the following contexts (this list is not exhaustive):

	• computer systems for organisations and businesses • ICT in business organisations • ICT as a marketing/advertising tool • ICT in tourism • ICT in the workplace.
The Performance (1) What skills or knowledge are best demonstrated through practical performance? Summed up in Outcomes 1 & 3 … Technology Process & Quality of Information Solutions	Productivity applications (IT skills). Communications using a network. Selection of systems. Application of design principles. Management of data, communications. e.g. Students create an advert. Students plan to start a business – organising office, information flows, databases, website etc. (this could be presented as a portfolio). **Applications and systems software** • introduction to common communication software for local area networks and connection of common office peripheral devices • commonly used applications software in organisations for productivity, planning and communication e.g. word processing, publishing, presentation and financial data management. **The nature, form and transfer of digital data** • digital documents and images as information in ICT solutions • characteristics of data for specific print or online formats • file formats, compression and encryption techniques, conversion, size and storage requirements • awareness of constraints in usage and communication over networks. **Processing and management of data** • efficient search, retrieval and referencing strategies • file and document management and archiving practices and procedures • input validation and manipulation of data (e.g. text-, numerical- and image-based), integration and presentation of these data • awareness of use of knowledge management systems for storage and retrieval. **Creative application of information design principles** • standards and conventions related to digital data and information for a specified target audience, purpose, platform and medium (e.g. online requirements) and form of communication • critical design standards and conventions • use of style sheets and automated functions • use of techniques for representing designs e.g. proposals, graphic outlines, storyboards, annotated diagrams, flowcharts, concept design or thumbnails. **Technology processes in the workplace** • problem-solving methodology incorporating ICT used within the organisation • research, planning and organisational skills • techniques to communicate plans and design e.g. graphic overviews and flowcharts • use of ICT to support enterprise, individual and collaborative interactions.
(2) What are the critical components of that practical performance?	Skills in using applications – including database applications. Creative application of principles, techniques and skills related to tasks. Demonstration of knowledge of concepts may only be seen in action

	e.g. change file formats. The digital product - determined by the task design (could be multimedia, database etc.).
(3) Why can't those components be demonstrated on paper?	Creative application of design principles and ICT skills – may need the audio-visual impact. Design of information solutions leads to digital products – without making the products it is difficult to evaluate the solution. Skills in using productivity applications is part of the required content of the course.
The Technologies (1) What alternative representations other than paper could be used?	In schools a range of ICTs are used but primarily these are focussed on desktop computers (peripheral devices – cameras). Students need some use of networks to develop knowledge of networking (e.g. using websites).
(2) What level of compromise in reliability, authentication and cost is acceptable in preference to NOT assessing the performance at all?	The expectations of students and parents (particularly interested in the area) are driving a requirement to adequately and authentically assess digital production and technical performance. Also leads to the investigation of new technologies. A constraint is equity of access to technologies at school to permit fair assessment of digital production and technical performance.
The Teachers (1) What are key characteristics of typical teachers? (2) How do these characteristics relate to assessment in the course? (3) How do these characteristics relate to the use of ICT in the course?	There is a wide range of teachers involved in the course in terms of ICT skills, knowledge, understanding of course, and previous involvement in teaching the course. This is exacerbated by a teacher shortage. It is likely that there are many teachers with little idea of standards expected of students and little experience in assessing the type of work students are required to complete. There is a critical issue of teacher workload and therefore to minimise the impact on teacher time the research should look for those with experience in BIT/IM. The research project should look at teachers involved in ARM Panels, working groups, Computer Fundamentals panels etc. Ian Gaynor is aware of about 30 suitable schools. Suggest that once the assessment task structure is in place we meet with them to generate ideas of the assessment task and call for volunteers to be involved.

Appendix C

Rationale and Analysis of Assessment Task

The overall form and structure for the assessment task(s) was based on a situation analysis.

Nature of Assessment Task(s)

The aim is to assess students completing Unit 2BAIT where the focus is information and communication technologies in business. Skills, principles and practices associated with various types of businesses to enhance students' career prospects are emphasised. Students examine the use of ICT in a range of administrative and business environments. They identify and explain the components and configuration of a computer system to meet the needs of the organisation. Students design information solutions for problems encountered in these contexts and understand the social issues inherent in work practices. Within the focus area of ICT in business, teachers may choose one or more of the following contexts (this list is not exhaustive):

• small commercial business
• large commercial business
• government organisation
• non-profit organisation

The situation analysis has determined that the following content underpinned by Outcomes 1 and 3 of the course is difficult, if not impossible, to assess using paper and pen exams but is conducive to digital forms of assessment.

Applications and systems software

• introduction to common communication software for local area networks and connection of common office peripheral devices
• commonly used applications software in organisations for productivity, planning and communication e.g. word processing, publishing, presentation and financial data management.

The nature, form and transfer of digital data

• digital documents and images as information in ICT solutions
• characteristics of data for specific print or online formats
• file formats, compression and encryption techniques, conversion, size and storage requirements
• awareness of constraints in usage and communication over networks.

Processing and management of data

• efficient search, retrieval and referencing strategies

- file and document management and archiving practices and procedures
- input validation and manipulation of data (e.g. text-, numerical- and image-based), integration and presentation of these data
- awareness of use of knowledge management systems for storage and retrieval.

Creative application of information design principles

- standards and conventions related to digital data and information for a specified target audience, purpose, platform and medium (e.g. online requirements) and form of communication
- critical design standards and conventions
- use of style sheets and automated functions
- use of techniques for representing designs e.g. proposals, graphic outlines, storyboards, annotated diagrams, flowcharts, concept design or thumbnails.

Technology processes in the workplace

- problem-solving methodology incorporating ICT used within the organisation
- research, planning and organisational skills
- techniques to communicate plans and design e.g. graphic overviews and flowcharts
- use of ICT to support enterprise, individual and collaborative interactions.

Analysis of Assessment Task

Task Component	**Outcomes**	**Content (from syllabus)**
Digital Portfolio – Product	Provides some confirmation of demonstration of O1 illustrated through Process Doc. Gives indications of the first two aspects of O3 that are referenced in Process Doc.	Applications and systems software • commonly used applications software in organisations for productivity, planning and communication e.g. word processing, publishing, presentation and financial data management. The nature, form and transfer of digital data • digital documents and images as information in ICT solutions • characteristics of data for specific print or online formats Processing and management of data • input validation and manipulation of data (e.g. text-, numerical- and image-based), integration and presentation of these data Creative application of information design principles • standards and conventions related to digital data and information for a specified target audience,

		purpose, platform and medium (e.g. online requirements) and form of communication • critical design standards and conventions • use of style sheets and automated functions
Digital Portfolio – Process Doc	All of O1 comprehensively addressed with confirmation in Product. Clearly addresses Aspect 3 of O3 but only partially addresses the other aspects in conjunction with Product.	The nature, form and transfer of digital data • characteristics of data for specific print or online formats Processing and management of data • efficient search, retrieval and referencing strategies • file and document management and archiving practices and procedures • awareness of use of knowledge management systems for storage and retrieval. Creative application of information design principles • standards and conventions related to digital data and information for a specified target audience, purpose, platform and medium (e.g. online requirements) and form of communication • critical design standards and conventions • use of techniques for representing designs e.g. proposals, graphic outlines, storyboards, annotated diagrams, flowcharts, concept design or thumbnails. Technology processes in the workplace • problem-solving methodology incorporating ICT used within the organisation • research, planning and organisational skills • techniques to communicate plans and design e.g. graphic overviews and flowcharts • use of ICT to support enterprise, individual and collaborative interactions.
Digital Portfolio – Extra Artefacts	Addresses "skills and techniques" in O3 and may provide indications of "forms, structures and conventions".	Applications and systems software • commonly used applications software in organisations for productivity, planning and communication e.g. word processing, publishing, presentation and financial data management. The nature, form and transfer of digital data • digital documents and images as information in ICT solutions • characteristics of data for specific print or online formats Processing and management of data • input validation and manipulation of data,

		integration and presentation of these data Creative application of information design principles • standards and conventions related to digital data and information for a specified target audience, purpose, platform and medium (e.g. online requirements) and form of communication • critical design standards and conventions • use of style sheets and automated functions
Exam – Reflective Questions	Addresses all of O1. May provide indications of “forms, structures and conventions” and “exploring alternatives” in O3.	Processing and management of data • efficient search, retrieval and referencing strategies • file and document management and archiving practices and procedures Technology processes in the workplace • problem-solving methodology incorporating ICT used within the organisation • research, planning and organisational skills • techniques to communicate plans and design e.g. graphic overviews and flowcharts • use of ICT to support enterprise, individual and collaborative interactions.
Exam – Performance Tasks	Addresses “skills and techniques” in O3 and may provide indications of “forms, structures and conventions”.	Applications and systems software • commonly used applications software in organisations for productivity, planning and communication e.g. word processing, publishing, presentation and financial data management. The nature, form and transfer of digital data • digital documents and images as information in ICT solutions • characteristics of data for specific print or online formats • file formats, compression and encryption techniques, conversion, size and storage requirements Creative application of information design principles • standards and conventions related to digital data and information for a specified target audience, purpose, platform and medium (e.g. online requirements) and form of communication • critical design standards and conventions • use of style sheets and automated functions

Appendix D

AIT Task Assessment: Assessment Criteria Portfolio

Component1 Portfolio Product -

Miss Shoppe Website

Develop a website that will allow her shop to have an online presence as a means of contacting her target audience, promoting her business and potentially selling more products. The website should include general information regarding the shop (Open hours, Products, Location), contact details (Location, Telephone number, Email address) and an online catalogue (List of products, Bulletin Board, Mailing List, Current News). Her corporate colours are Green, White and Black.

Criterion 1 Application of design principles through styles/structures/codes/conventions relevant to the form of information solution (e.g. Digital Graphics - spacing, colour, size, position, fonts, consistency, variations ...)

Application of web-site design principles for shop

Consistency of layout or operation or interaction.

Left-to-right and Top-to-bottom layout

Use of graphics, text, colours and effects

Criterion 2 Application of technical skills and techniques (e.g. Digital Graphics - layers, tools, file format, editing)

Placing of graphics and text

Links from text, buttons, graphics

Features such as roll-overs

[Should have an operational example of each feature but does not need to be completed]

Criterion 3 Creativity and innovation in design to produce functional and aesthetic features

Creative use of graphics and other features

Criterion 4 Appropriateness of product design for the likely target audience and/or client

Suited to target audience?

Meets requirements of client - general info, contact details, online catalogue, and corporate colours.

Component 2 Process Document

The Design Process Document - Shop Website

A website to allow a shop to contact target audience, promote business and sell more products.

Collated in five hours with a maximum of NINE pages as a single PDF file that comprises four sections: (1) Investigate (6%) - 2 pages, (2) Design (4%) - 3 pages, (3) Production (10%) - 3 pages, (4) Evaluate (10%) - 1 page

Criterion 1 Describes the human need or opportunity.
Defines the main objectives of the information solution. Purpose, meaning, audience and client specifications. Alternative solutions - compare aspects.
Strategies used to find and analyse information.
Criteria used to evaluate the success of own solution.

Criterion 2 Information that would allow another skilled person to complete the production.
One design example that was improved, and why.
Explain creative application of technologies.
Design elements, standards and conventions used.

Criterion 3 Production plan with the amounts of time for each stage.
Hardware, software, and other materials/people involved.
Skills that were needed to use the hardware and software.

Criterion 4 Evaluation criteria used.
Results of evaluation explaining - strengths/weaknesses.
Future improvements to the technology process.

Component 3 Two Artefacts
Two digital artefacts should be submitted that illustrate design and development skills in any two of the following domains … graphics, databases, spreadsheets, web-publishing etc. For each artefact, complete the table provided in no more than ONE page in length.

Criterion 1 Breadth of types of software/hardware, skills and processes
Look at support document. Must illustrate design and development skills in two different domains … graphics, databases, spreadsheets, web-publishing etc.
Must be different skill areas from their portfolio product.
Need to get a feel for their breadth and depth of skills. Looking for generic skill development those impacts on products.

Criterion 2 Creative use of software and hardware
Look at the two digital artefacts.
Creativity is not only visual - the design of a database or spreadsheet may creatively use the features available to enhance the performance or user experience.

Criterion 3 Selection of software and hardware
Look at support document and artefacts.
Appropriateness of software and hardware to the tasks, purposes, audience etc.

May be simple non-commercial software.

Criterion 4 Selection of techniques
Look at support document and artefacts.
May need to interpret what techniques were used from the operation/view of the resulting artefact.

Criterion 5 Application of skills
Look at support document and artefacts.
May need to interpret what skills were used from the techniques evident in the resulting artefact.

Appendix E

AIT Task Assessment: Assessment Criteria Examination

Component 4 Exam Section A

Criterion 1 Explanation of investigation of information solutions (Question 1 - For the project you presented in your portfolio think about how you investigated and evaluated information solutions.)

(a) (a) What was the technology challenge you were responding to?

(b) What objectives did you develop for the project?

(c) What research did you undertake and how did this impact on design decisions?

(d) What were the TWO most critical factors in the selection of your final design proposal?

Criterion 2 Explanation of design and planning processes (Question 2 - For the project explain, with appropriate justification, how you devised, communicated and evaluated proposals and design plans.)

Criterion 3 Explanation of production processes and management of data (Question 3 - Production processes are required to make a prototype or final product from your designs.)

(a) Explain the main production processes for your information solution.

(b) What modifications did you have to make?

(c) How did you ensure you used time efficiently and resources safely and effectively?

Criterion 4 Explanation of evaluation of production processes (Question 4)

(a) Describe how you evaluated your information solution.

(b) Explain the standards and conventions that were appropriate for your information solution.

(c) Describe the changes you would have liked to make to your information solutions explaining how this would improve the quality of your solution.

Component 5 Exam Section B (Practical)

Criterion 1 Creative exploration of design options of the relevant forms, structures and conventions.(Task 1)

Develop two different designs for a resort logo and two different designs for an A4 tri-fold brochure to advertise the resort. Add notes to your designs to explain the creative processes involved in developing the logo and brochure.

The designs are the first part of the main PDF file either created on computer or scanned from hand drawn designs.

Criterion 2 Logo (Tasks 2, 5 & 6)

Create a logo for the resort that will be used in your A4 tri-fold brochure - Save your logo as **logo.jpg**

Will need to consider reflection (oral and PDF) to give a 4.
The logo should be part of the brochure that is in the main PDF file. If necessary the logo can be accessed separately as a graphic.

Criterion 3 Spreadsheet graphs (Task 3)
Import the climatic data into a spreadsheet and create at least two different graphs using the data. Save the spreadsheet file with the graphs included as graphs.xls
The graphs should be part of the brochure that is in the main PDF file. If necessary the spreadsheet can be accessed separately as an Excel file.

Criterion 4 Brochure (Task 4, 5 & 6)
Using some of the digital photographs supplied, the data found on the file data.txt, and your own ideas develop a tri-fold brochure to promote the hotel. The brochure must include
• logo, the address and contact details of the resort
• financial information about the cost of staying at the resort
• climatic information (temperature and rainfall) at the resort presented graphically
• the text that introduces and sells the resort, explains or captions the images you have chosen and holds the brochure together.
Save the brochure as brochure.pdf

Criterion 5 Styles, forms, structures, conventions and techniques used with digital data. File formats (Submission)

Consider overall design of products and reflection in audio file.
logo.pdf, brochure.pdf, brochure.mp3 (or brochure.wav, reflection.doc
Will need to consider reflection (oral and PDF) to give a 3. Students typed their reflections that should be part of the main PDF file. They should have made an audio recording of their reflections but in two schools that was not possible and for a few individual students errors were made in saving and thus the audio file is not available.
For File Formats (Submission) refer to technicians notes for omissions otherwise assume files are OK.

Appendix F

Assessment Task: Portfolio

Product

<<NAME OF PROJECT>> (20 marks)

Your design brief is as follows, you have 15 hours of class time over 4 weeks to investigate, design, produce and evaluate a prototype of an information solution that will run on a computer in the classroom.

<<INSERT OWN DESIGN BRIEF>>

Example Design Brief

Miss Shoppe is the manager at a local retail clothes outlet. She is very concerned with the increasing number of people shopping online and the declining number of consumers venturing into her shop to purchase her products. The shops target market is teens (12 – 20 years). She has approached you to create her own online shop front. She would like the website to include general information regarding the shop (Open hours, Products, Location), contact details (Location, Telephone number, Email address) and an online catalogue (List of products, Bulletin Board, Mailing List, Current News). Her corporate colours are Green, White and Black.

Using this information, design the online presence for Miss Shoppe. Miss Shoppe has requested that you present your designs as detailed storyboards and provide a summary of recommendations that you have made. Miss Shoppe has also requested that a detailed production plan be developed.

Select your best design and develop a website that will allow her shop to have an online presence as a means of contacting her target audience, promoting her business and potentially selling more products. Use any suitable software to create the website and any suitable media, taking care to appropriately acknowledge the source of any media you use.

Your digital product should:

- suit the intended purpose and audience/users;
- meet the requirements of the design brief and/or client specifications;
- illustrate creative application of information design principles and technologies;
- make use of appropriate information structures, forms, layouts and symbols;

- employ relevant standards and conventions to create multi-tiered information solutions; and
- use appropriate methods and techniques to represent the design of information solutions.

You must only use computer hardware and software provided by the school and must acknowledge any help that you receive from other people and sources of information.

Follow a technology process to investigate, design, produce and evaluate your product. Output from these processes will be required for the Design Process Document you will be asked to collate when you have finished. Keep all your electronic documents in a folder and paper documents in a document wallet.

6 marks: Application of design principles 6 marks: Application of technical skills and techniques 4 marks: Creativity and innovation in design to produce functional and aesthetic features 4 marks: Appropriateness of product design for the likely target audience and/or client

Your Technology Process

(1) INVESTIGATE

Investigate possible solutions for the design brief, including the following:

- Define in your own words the human need or opportunity that is given in the design brief.
- Define the main objectives of the information solution you will need to produce.
- Quickly search for information on at least two existing or similar solutions.
- For each solution briefly describe what aspects you like and don't like.
- Use all your information to generate ideas for your own solution. You could use methods such as brainstorming and mind-mapping.
- List and describe the criteria that you will use to evaluate the success of your own solution.

(2) DESIGN

Develop a design for your information solution and present it using appropriate forms such as: descriptions, storyboards, thumbnail sketches, annotations, photographs, drawings, flowcharts and schematics. In developing your design take note of the following points.

- Provide information that would allow another skilled person to complete the production of your solution.

- Make sure you keep examples of all designs even if you did not end up using them.
- Aim to apply technologies in creative and original ways to meet the need.
- Always keep in mind the purpose, meaning, target audience and client specifications.
- Explain the design elements, standards and conventions you have used in your design.

(3) PRODUCTION

Develop a plan of production and then make a prototype product of your solution. Your plan of production should include:

- A production plan (e.g. timeline, Gantt chart) with estimated amounts of time.
- Describe the hardware, software, and any other materials or people involved.
- A list of the skills that will be needed to use the hardware and software to make your solution.

You may use any of the software available on the school's computers for productivity, planning and communication (e.g. word processing, publishing, presentation and data management).

Your prototype product needs to be delivered in a single digital file with one of the following formats: PDF, AVI, JPG, GIF, SWF, FLA, HTML or ZIP (must be a collection of files with the permitted formats e.g. zipped folder of a website of HTML and FLA files). The file should not exceed 20MB. Name your file as **<Your ID>product.xxx**

(4) EVALUATION

Evaluate your prototype information solution and technology processes, including:

- A list of evaluation criteria, each described in one or two sentences.
- A summary of the results of your evaluation explaining the strengths and weaknesses of your solution.
- A list of improvements you would make of the technology process you used, each described in one or two sentences.

The Design Process Document Template

(30 marks - total)

Student Name:

Student ID:

Teacher Name:

Date Submitted:

You have five hours to collate a document with a maximum of NINE pages that comprises four sections: Investigate Design, Production and Evaluation. Relevant language and terminology should be used throughout.

Use this MS Word document as a template by putting text and/or graphics into the boxes provided. When you have finished print it as a PDF file to hand in. Name your file as **<Your ID>process.pdf**

Use the information you collected while you were completing your project, select the most useful information and add any explanations or annotations. You may include diagrams, photos, screenshots etc.

For each section there is ‘common’ required content and ‘creative’ individual choice content.

(1) INVESTIGATE (6 marks)

In no more than TWO pages present the results of your investigation of solutions to the information problem.

Required Content

Describe the human need or opportunity that was addressed.

Define the main objectives of the information solution. Include the purpose, meaning, target audience and client specifications.

Describe TWO alternative solutions you considered in your investigation and explain what aspects you liked and didn’t like.

A summary of the strategies that were used to find and analyse relevant information to generate ideas including methods such as brainstorming and mind-mapping.

List and describe the criteria that you used to evaluate the success of your own solution.

Your Choice

Put any additional information that you believe is important but does not fit the required content section. [Remove this message]

(2) DESIGN (4 marks)

In no more than THREE pages present your final design and design processes.

Required Content

Provide information that would allow another skilled person to complete the production such as descriptions, storyboards, thumbnail sketches, annotations, photographs, drawings, flowcharts and schematics developed to represent the design.

- Give one design example that you improved, explaining what improvements you made and why.
- Explain how you applied technologies in creative and original ways to meet the need.
- Explain the design elements, standards and conventions you used in your design.

Your Choice

Put any additional information that you believe is important but does not fit the required content section. [Remove this message]

(3) PRODUCTION (10 marks)

Document your production in no more than THREE pages.

Required Content

- Provide your production plan with the amounts of time for each stage.
- Describe the hardware, software, and any other materials or people involved.
- List the skills that were needed to use the hardware and software to make your solution.

Your Choice

Put any additional information that you believe is important but does not fit the required content section. [Remove this message]

(4) EVALUATION (10 marks)

In no more than ONE page present the evaluation of the prototype information solution and technology processes employed.

Required Content

- List the evaluation criteria you used. Describe each in one or two sentences.
- Summarise the results of your evaluation explaining the strengths and weaknesses of your solution.
- List improvements you would make to the technology process you used. Describe each improvement in one or two sentences.

Your Choice

Put any additional information that you believe is important but does not fit the required content section. [Remove this message]

Two Extra Digital Artefacts

Two digital artefacts should be submitted that illustrate your design and development skills in any two of the following domains … graphics, databases, spreadsheets, web-publishing etc. You must have created the digital artefacts at school, under supervision from your teacher. Any assistance from the teacher or others must be explained. For each artefact, complete the table provided in no more than ONE page in length.

Each of your digital artefacts is to be delivered in a single digital file with one of the following formats: PDF, AVI, JPG, GIF, SWF, FLA, HTML or ZIP (must be a collection of files with the permitted formats e.g. zipped folder of a website of HTML and FLA files). Each file should not exceed 20MB. Name your files as **<Your ID>artefact1.xxx** and **<Your ID>artefact2.xxx**

The Digital Artefacts Description Template

(20 marks: 10 marks per artefact)

Student Name:

Student ID:

Teacher Name:

Date Submitted:

In ONE page describe for each artefact what hardware, software, techniques and skills were needed.

Digital Artefact One

Description of Artefact	
Hardware Used	
Software Used	
Techniques Used	
Skills Used	
Help from others	
Other explanations	

Digital Artefact Two

Description of Artefact	
Hardware Used	
Software Used	
Techniques Used	
Skills Used	
Help from others	
Other explanations	

Appendix G

Assessment Task: Examination

Section A: Reflective Questions (1 hour) (10 marks)

Type in the boxes your responses to the questions.

You may view your portfolio product.

Student ID:

1. For the project you presented in your portfolio think about how you investigated and evaluated information solutions. (3 marks)

(a) What was the technology challenge you were responding to?

(b) What objectives did you develop for the project?

(c) What research did you undertake and how did this impact on design decisions?

(d) What were the TWO most critical factors in the selection of your final design proposal?

2. For the project explain, with appropriate justification, how you devised, communicated and evaluated proposals and design plans. (2 marks)

(a) Explain how you went about devising a design/s for an information solution.

(b) Explain the forms and techniques you used to communicate your design/s.

(c) Explain how you evaluated your proposals and design plans.

3. Production processes are required to make a prototype or final product from your designs. (3 marks)

(a) Explain the main production processes for your information solution.

(b) What modifications did you have to make?

(c) How did you ensure you used time efficiently and resources safely and effectively?

4. For your project, consider the evaluation of your information solution and your production processes. (2 marks)

(a) Describe how you evaluated your information solution.

(b) Explain the standards and conventions that were appropriate for your information solution

(c) Describe the changes you would have liked to make to your information solutions explaining how this would improve the quality of your solution.

Section B: Performance Tasks (2 hours) (20 marks)

The Challenge

Ningaloo Reef Resort is opening a 200 room hotel and marina on the coast of Western Australia. The resort is in a sensitive conservation area and there was a lot of opposition to the development. The hotel itself offers fine dining, luxurious rooms, swimming pools as well as swimming, diving and snorkelling on the reef. It is the intention of the owners that the resort is presented as having a low environmental impact with a ‘clean and green' image. The hotel manager wants you to apply your information technology skills to a series of tasks to help with the marketing of the resort. Your tasks are as follows:

Task 1: Planning Logo and Brochure (required time 15 minutes) (5 marks)

You are provided with a template, you may either work on the paper provided or using the file plan_template.doc or plan_template.ppt. You will not be permitted to continue into Task 2 until the 15 minutes has elapsed.

Develop **two** different designs for a resort logo and **two** different designs for an A4 tri-fold brochure to advertise the resort. Add notes to your designs to explain the creative processes involved in developing the logo and brochure. To help you, think about your ideas, themes, design principles, colours, shapes and symbolism. Make any notes clearly as the designs will be scanned for assessment. Put your candidate number at the top of each page.

Task 2: Logo (suggested time 30 minutes) (3 marks)

Take **one** of your design ideas and using **any** software available to you, create a logo for the resort that will be used in your A4 tri-fold brochure. Save your logo as **logo.jpg** on the mass storage device provided.

Task 3: Graphs (suggested time 15 minutes) (2 marks)

In the file data.txt there is some climatic data. In the next task you will be creating and brochure in which you will be asked to present the climatic data graphically. Import the data into a spreadsheet and create at least two different graphs using the climatic data. Save the spreadsheet file with the graphs included as **graphs.xls** on the mass storage device provided.

Task 4: Brochure (suggested time 45 minutes) (3 marks)

Using some of the digital photographs supplied, the data found on the file **data.txt**, and your own ideas develop a tri-fold brochure to promote the hotel. The brochure is designed for travel agents and tourist offices. You may manipulate the photographs in any way you wish but **only** these files may be used. The brochure **must** include

- your **logo**, the **address** and **contact** details of the resort
- **financial information** about the cost of staying at the resort
- **climatic** information (temperature and rainfall) at the resort presented **graphically**
- the text that introduces and sells the resort, explains or captions the images you have chosen and holds the brochure together. How much you write is up to you.

Save the brochure as **brochure.pdf** on the mass storage device provided.

Task 5: Prepare Reflection (suggested time 10 minutes) (3 marks)

Prepare a 30 second (half a minute) audio reflection by listing headings or points in the document **reflection.doc**

- explain the creative processes involved in developing the logo and brochure
- appraise your finished product
- other uses and other products that could be developed

Think about your ideas, themes, the images chosen, design principles, colours, shapes and symbolism of your logo and brochure as well as what you achieved and would have liked to achieve. Open and make ready a suitable sound recording application but **do not begin** recording until the invigilator gives permission.

Task 6: Audio recording (required time 5 minutes) (2 marks)

The invigilator will announce that audio recording may begin.

Record the audio commentary.

Save the audio file as **reflection.mp3** or **reflection.wav**

Submission checklist (2 marks for submission of correct file formats)

1) **Design ideas for logo and brochure**, sketches and notes (on paper or plan_template.doc)
2) **logo.pdf** (saved to the mass storage device provided).
3) **brochure.pdf** (saved to the mass storage device provided).
4) **brochure.mp3 (or brochure.wav)** (saved to the USB Flash storage device provided).
5) **reflection.doc** (saved to the mass storage device provided).

Appendix H

Student Survey

APPLIED INFORMATION TECHNOLOGY

This survey is part of research being conducted by Edith Cowan University (ECU).

Please read the following Disclosure Statement carefully as it explains what this research is about.

Disclosure Statement

This questionnaire forms part of the evaluation of the use of computers at the school to help the assessment of learning. The evaluation is being conducted by a team from ECU and is designed to provide your school with the most accurate information possible. What you as a student think and the activities you are involved in at school are very important to this evaluation and therefore we are surveying students from your class to collect this information.

Your responses will be strictly confidential, only the ECU evaluation team members will see your particular responses. The information will be collated with no reference to individuals and no identifying information for reports to the school and teachers at the school. Such reports will only include general and summary information and will in no manner identify individual or groups of students or teachers.

Instructions to Students

Please do not write your name on the survey sheet. **Put your ID code on the sheet**, only this will be recorded and known only to the research team. The ID code will maintain the confidentiality of your responses and also provide a way of re-identifying your data if you choose to withdraw from the project.

To ensure maximum confidentiality all the questionnaires from your class will be placed in a sealed envelope to be returned to Edith Cowan University. Therefore no one at your school will see your questionnaire.

It should take you about 15 to 20 minutes to answer the questions but take as long as you need.

Please use PENCIL so that you can erase and change responses if necessary.

Some items require you to CIRCLE or TICK an alternative while others provide the opportunity for you to write brief responses (note form is OK).

CAREFULLY ANSWER THE QUESTIONS ON THE FOLLOWING PAGES

Student Survey – Digital Assessment Project

Please circle ONE response for each row.

Gender (circle): Male / Female

Doing exams in the computer laboratory

E1.	(a)	How often have you done an exam or test on a computer before?	**Lots**	**Some**	**Little**	**None**
	(b)	How much more time would you need to get used to it?	**Lots**	**Some**	**Little**	**None**

Doing the Applied Information Technology exam

E2.	(a)	It was easy to use the computer for doing the exam.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(b)	It was easy to use the computer in the exam to develop my design ideas.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(c)	The computer was a quick way for presenting my design ideas in the exam.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(d)	The computer was good to create my logo, graphs and brochure in the exam.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(e)	The computer was good for reflecting on my design ideas in the exam.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(f)	The computer was good for completing the questions in	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**

SECTION A in the exam.

(g)	It was easy to follow the steps of the exam on the computer.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
(h)	The steps of the exam helped me to develop my design ideas.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
(i)	Overall, the computer is a good tool for designing products in an exam.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
(j)	Overall, I was able to show what I can do in the exam.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
(k)	Overall, it was better doing the exam using a computer than on paper.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**

E3. The two best things about doing the Applied Information Technology exam in the computer laboratory:

👍

👍

E4. The two worst things about doing the Applied Information Technology exam in the computer laboratory:

👎

👎

This section asks questions about your digital portfolio that included a product you made in a project, a process document and two extra artefacts.

Doing portfolios using computers

P1	(a)	How often have you done a portfolio on a computer before?	**Lots**	**Some**	**Little**	**None**
	(b)	How much more time would you need to get used to it?	**Lots**	**Some**	**Little**	**None**

Doing the Applied Information Technology portfolio

P2.	(a)	It was easy to use the computer for doing the portfolio.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(b)	It was easy to use the computer for my portfolio to develop my ideas.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(c)	The computer was a quick way for presenting my ideas in the portfolio.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(d)	The computer was good to create my product for the portfolio.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(e)	The computer was good for reflecting on my ideas in the process document for the portfolio.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(f)	The computer was good for showing my skills in the portfolio in the product and	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(g)	It was easy to follow the steps to create the portfolio on the computer.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(h)	The steps of the portfolio helped me to develop my ideas.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(i)	Overall, the computer is a good tool for creating portfolios.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(j)	Overall, I was able to show what I can do in the portfolio.	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
	(k)	Overall, it was better doing the portfolio using a computer than	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**

P3. The two best things about doing the Applied Information Technology portfolio:

-
-

P4. The two worst things about doing the Applied Information Technology portfolio:

-
-

Experience and Knowledge with Computer Technology

5. What do you have at **home** (circle ANY of the following that apply to you)

Computer **Colour Printer** **Digital Camera** **Video Camera** **MP3 Player**
(e.g. iPod) **Mobile Phone** **Laptop** **Game Console DVD Player**

DVD Burner

6. Do you have Internet access at **home** (circle ONE)

NO Internet **Dial-up Internet** **Broadband Internet**

7. Circle the response that best describes how often you use a computer at home.

Most Days **More than once a week** **Most Weeks** **Rarely**

8. Estimate the amount of time in MINUTES you spent using computers at school on each day LAST WEEK.

Monday	Tuesday	Wednesday	Thursday	Friday

9. When you type do you try to touch type (use all of your fingers)? **YES** or **NO**

10. Do you, or would you, use a computer to do the following tasks? (Circle ONE for each)

(a) Keep a list of telephone numbers and addresses of friends. **I do** **I would** **No**

(b) Draw a diagram or picture. **I do** **I would** **No**

(c) Type an assignment for school. **I do** **I would** **No**

(d) Do a line graph or pie-diagram as part of an assignment. **I do** **I would** **No**

(e) Send a letter to every club member or friend **I do** **I would** **No**

11. Circle YES or SOMETIMES or NO to show whether you agree with each of the following statements.

a) Using computers makes the work **at school** more difficult.

YES or SOMETIMES or NO

b) I enjoy using computers **at school**.

YES or SOMETIMES or NO

c) I like to use a computer **at home** to do school work.

YES or SOMETIMES or NO

d) I like to find things out for myself instead of being told by the teacher.

YES or SOMETIMES or NO

e) Computers are good for the world

YES or SOMETIMES or NO

12. Circle either “YES”, “Not Sure” or “NO”.

(a)	I feel confident working with computers.	**YES**	**Not Sure**	**NO**
(b)	I'm good at using computers. . . . .… .	**YES**	**Not Sure**	**NO**
(c)	I feel OK about trying a new problem on the computer.	**YES**	**Not Sure**	**NO**
(d)	I usually do well with computers. .	**YES**	**Not Sure**	**NO**
(e)	I could learn to program a computer. .	**YES**	**Not Sure**	**NO**
(f)	Using a computer is very hard for me. .	**YES**	**Not Sure**	**NO**

13. Rate yourself on your skill level in using each of these types of computer software and equipment. For each row TICK the CELL that best describes your skills.

a	**Word processor**	I can’t do	I can print a	I can insert images,	I can use columns

		much	document, change fonts, spell check, and insert a footer and page numbers.	create tables, change Page Setup, and change margins.	and sections, set up styles, and use mail merge.
b	**Spreadsheets**	I can't do much	I can enter data, use Sort, create charts [graphs] and modify them.	I can insert some calculations, format cells, insert and delete rows and columns.	I can use complex formulae, use absolute and relative cell references.
c	**Databases**	I can't do much	I can create data files, enter data, and use simple queries to retrieve data.	I can create simple tables, use wizards to create reports and forms.	I can create a relational database.
d	**Slideshow software (e.g. PowerPoint)**	I can't do much	I can create a slideshow, insert images, change font and layout.	I can navigate during a presentation, add animation and transitions, insert hyperlinks.	I can create a master slide, include sound, print hand-outs, add navigation buttons.
e	**Email**	I can't do much	I can send and access emails, and add to and access Address book entries.	I can store messages in folders, locate Sent and Deleted messages, manage the Address book.	I can add a Signature, and add attachments.
f	**Computer File Management**	I can't do much	I can save files in a folder, create and name folders, navigate between folders, copy, delete and rename files.	I can recognise different file types, navigate between Drives and Directories, access a network, use Help files.	I can zip and unzip files, install software.
g	**The Internet**	I can't do much	I can navigate to known web sites, create Favourites, do basic searches.	I can save images and text, use Advanced search tools, organise Favourites.	I can conduct complex searches, download and install plugins, use different browsers, alter browser preferences.

h	**Web page authoring**	I can't do much	I can create pages and links, insert and format text, insert images.	I can use tables, create external links and email links.	I can create a website with pages and folders, insert sound, upload files to the web.
I	**Digital photography**	I can't do much	I can take photos or video, transfer to a computer.	I can review images/video on camera, adjust camera settings such as flash and close-up.	I can adjust camera menu options such as resolution.
j	**Image editing**	I can't do much	I can do simple editing such as crop, delete and draw.	I can change image size, format and resolution.	I can undertake complex image manipulation using filters and other special effects.
k	**Video editing**	I can't do much	I can do simple editing such as crop, delete and insert.	I can use basic software to introduce transitions, import and edit sound track, add titles and subtitles.	I can use advanced software to apply complex editing and special effects.

THANK YOU FOR YOUR HELP.

Appendix I

Student Forum

Digital Forms of Assessment

School ………………………………. Date ……………………………….

Looking back on the **AIT portfolio and practical exam** that you did a few weeks ago, we would like your thoughts to be part of our research report. Your comments will be attributed anonymously as a group (e.g.) as ‘student group 6’.

What did **you** think of the task(s) you were asked to do?

What were the **reactions of other students’** to the task(s)?

Did the **computers help**?

How much different was this to **how it used to be done**?

What, if anything, would you like **changed** in future?

Were there any **technical problems** with doing the activities?

Were there any **other problems** with the activities?

Any other thoughts … or suggestions for developing the use of digital forms of assessment?

We are really very grateful for your help.

Appendix J

Teacher Interview

Digital Forms of Assessment Name ……………………………….

Looking back on the **AIT Portfolio and Exam** that you ran with your students this year, we would like your thoughts to be part of our research report. Your comments will be attributed anonymously (e.g.) as ‘teacher 6’, and we would like to use any quotes that help us to capture the event, the atmosphere of the activities, and your thoughts about it. Add or delete dot-points as required.

What did **you** think of the assessment tasks overall?

•

What did **you** think of the structure of the activities? (timing / sub-tasks / instructions)

•

What were **the students’** reactions to the activities?

•

What do **you** think of its potential? (for AIT or other subjects)

•

What did **you** think of the **quality** of work produced by your students for these tasks?

•

Were you **surprised** by the performance/confidence of any students? (pleased / disappointed)

•

What was the general **feedback** from students? (would they like more of it?)

•

Were there any **technical problems** with implementing the activities?

•

Were there any **other problems** with implementing the activities?

•

Any other thoughts … or suggestions for developing the use of digital forms of assessment?

•

We are really very grateful for your help in completing this form.

Appendix K

Survey data from case studies

CA Descriptive Statistics of Students' Questionnaire Responses (N=20

(N=20)	**Abbreviated Questions**	**Min**	**Max**	**Pop Mean**	**Mean**
Year	Student's schools year	11	11	11.13	11.00
Sex		Male=1	Female=2	1.20	1.31
q1a	How often have you done an exam on a computer before?	Often =1	None =4	3.20	3.68
q1b	How much more time would you need to get used to it?	Lots =1	None =4	2.80	2.64
q2a	It was easy to use the computer for doing the exam.	SA=1	SD=4	1.83	2.14
q2b	It was easy to use the comp. in the exam to develop ideas.	SA =1	SD =4	2.14	2.45
q2c	The computer was a quick way for presenting ideas in exam	SA =1	SD =4	1.96	2.05
q2d	The computer was good to create logo, graphs, brochure.	SA =1	SD =4	1.41	1.55
q2e	The computer was good for reflecting on my design ideas	SA =1	SD =4	1.85	2.18
q2f	The computer was good for completing Sect A in the exam.	SA =1	SD =4	1.65	1.86
q2g	It was easy to follow the steps of the exam on the computer.	SA =1	SD =4	1.78	2.05
q2h	The steps of the exam helped me to develop my ideas.	SA =1	SD =4	1.92	2.18
q2i	The computer is a good tool for designing products exams.	SA =1	SD =4	1.73	2.00
q2j	Overall, I was able to show what I can do in the exam.	SA =1	SD =4	1.90	2.14
q2k	It was better doing the exam on a computer than on paper.	SA =1	SD =4	1.68	2.00
p1a	How often have you done a portfolio on a computer before?	Lots=1	None=4	2.47	2.95
p1b	How much more time would you need to get used to it?	Lots=1	None=4	2.67	2.41
p2a	It was easy to use the computer for doing the portfolio	SA =1	SD =4	1.89	2.18
p2b	It was easy to use the computer for my portfolio for ideas	SA =1	SD =4	2.05	2.27
p2c	The computer was a quick way for presenting design ideas	SA =1	SD =4	1.87	2.23
p2d	The computer was good to create my portfolio product	SA =1	SD =4	1.68	1.86
p2e	The computer was good for reflecting in the process doc	SA =1	SD =4	1.94	2.14
p2f	The computer was good for showing my skills in the portfolio	SA =1	SD =4	1.76	2.09
p2g	It was easy to follow the steps to create the portfolio	SA =1	SD =4	1.97	2.14
p2h	The steps of the portfolio helped me to develop my ideas	SA =1	SD =4	2.00	2.09
p2i	Overall, the computer is a good for portfolios	SA =1	SD =4	1.64	2.05
p2j	Overall, I was able to show what I can do in the portfolio.	SA =1	SD =4	1.86	2.32
p2k	It was better doing the portfolio on a computer than on paper	SA =1	SD =4	1.58	2.18
q5_com	Which of the following do you use at home? Computer	No=0	Yes=1	.89	1.00
q5_dig	Digital camera	No=0	Yes=1	.85	.95
q5_vid	Vide0 camera	No=0	Yes=1	.63	.86
q5_mp3	MP3 player	No=0	Yes=1	.95	1.00
q5_lap	Laptop	No=0	Yes=1	.66	.77
q5_gam	Game console	No=0	Yes=1	.82	.86
q5_mob	Mobile phone	No=0	Yes=1	.93	1.00
q6	Do you have Internet access at home?	Dial up=2	Bdband=3	2.92	3.00
q7	How often do you use a computer at home?	Daily=1	Weekly=3	1.14	1.18
q8mon	Time in minutes spent using computers at school last week on: Monday	0	120	97.57	72.05
q8tue	Tuesday	0	120	94.92	74.64
q8wed	Wednesday	0	180	98.64	76.91
q8thu	Thursday	0	120	85.36	47.73

q8fri	Friday	0	180	103.21	85.95
q9	Touch type?	Yes=1	No=2	1.27	1.18
q10a	Keep list of telephone numbers/addresses	I do=1	No=3	2.28	2.41
q10b	Draw a diagram or picture.	I do=1	No=3	2.01	2.14
q10c	Type an assignment for school	I do=1	No=3	1.10	1.23
q10d	Do a line graph or pie-diagram as part of an assignment.	I do=1	No=3	1.42	1.59
q10e	Send a letter to every club member or friend.	I do=1	No=3	1.54	1.64
q11a	Using computers makes the work at school more difficult	Yes=1	No=3	2.70	2.59
q11b	I enjoy using computers at school.	Yes=1	No=3	1.30	1.27
q11c	I like to use a computer at home to do school work	Yes=1	No=3	1.26	1.32
q11d	I like to find things out for myself instead of being told.	Yes=1	No=3	1.75	1.86
q11e	Computers are good for the world	Yes=1	No=3	1.43	1.50
q12a	I feel confident working with computers	Yes=1	No=3	1.13	1.18
q12b	I'm good at using computers	Yes=1	No=3	1.30	1.23
q12c	I feel OK about trying a new problem on the computer	Yes=1	No=3	1.30	1.50
q12d	I usually do well with computers.	Yes=1	No=3	1.23	1.41
q12e	I could learn to program a computer.	Yes=1	No=3	1.71	1.68
q12f	Using a computer is very hard for me	Yes=1	No=3	2.90	2.86
q13_wp	Word processing: scale 1 to 4 low ability =1 high ability =4	1	4	3.60	3.27
q13_ss	Spreadsheet: scale 1 to 4 low ability =1 high ability =4	1	4	2.98	2.50
q13_db	Database: scale 1 to 4 low ability =1 high ability =4	1	4	2.65	2.59
q13_sl	Slideshow: scale 1 to 4 low ability =1 high ability =4	1	4	3.53	3.45
q13_em	Email: scale 1 to 4 low ability =1 high ability =4	1	4	3.63	3.77
q13_fm	File Management: scale 1 to 4 low ability =1 high ability =4	1	4	3.39	3.23
q13_in	Internet Research: scale 1 to 4 low ability =1 high ability =4	1	4	3.73	3.73
q13_wa	Web Authoring: scale 1 to 4 low ability =1 high ability =4	1	4	3.20	2.77
q13_dp	Digital Photography: scale 1 to 4 low ability =1 high ability =4	1	4	3.49	3.41
q13_ie	Image Editing: scale 1 to 4 low ability =1 high ability =4	1	4	3.54	3.14
q13_dv	Digital Video: scale 1 to 4 low ability =1 high ability =4	1	4	2.90	2.77

(SA=Strongly Agree, SD=Strongly Disagree)

LA Descriptive Statistics of Students' Questionnaire Responses (N=10)

(N=10)	**Abbreviated Questions**	**Min**	**Max**	**Mean**	**Pop. Mean**
Year	Student's schools year	11	11	11.00	11.13
Gender		Male=1	Female=2	1.17	1.20
q1a	How often have you done an exam on a computer before?	Often =1	None =4	3.00	3.20
q1b	How much more time would you need to get used to it?	Lots =1	None =4	3.11	2.80
q2a	It was easy to use the computer for doing the exam.	SA=1	SD=4	1.78	1.83
q2b	It was easy to use the comp. in the exam to develop ideas.	SA =1	SD =4	2.11	2.14
q2c	The computer was a quick way for presenting ideas in exam	SA =1	SD =4	2.11	1.96
q2d	The computer was good to create my logo, graphs, brochure.	SA =1	SD =4	1.56	1.41
q2e	The computer was good for reflecting on my design ideas	SA =1	SD =4	1.78	1.85
q2f	The computer was good for completing Sect A in the exam.	SA =1	SD =4	1.56	1.65
q2g	It was easy to follow the steps of the exam on the computer.	SA =1	SD =4	1.67	1.78
q2h	The steps of the exam helped me to develop my ideas.	SA =1	SD =4	2.00	1.92
q2i	The computer is a good tool for designing products exams.	SA =1	SD =4	1.78	1.73
q2j	Overall, I was able to show what I can do in the exam.	SA =1	SD =4	1.67	1.90
q2k	It was better doing the exam on a computer than on paper.	SA =1	SD =4	1.78	1.68

p1a	How often have you done a portfolio on a computer before?	Lots=1	None=4	3.00	2.47
p1b	How much more time would you need to get used to it?	Lots=1	None=4	2.89	2.67
p2a	It was easy to use the computer for doing the portfolio	SA =1	SD =4	1.78	1.89
p2b	It was easy to use the computer for my portfolio for ideas	SA =1	SD =4	2.11	2.05
p2c	The computer was a quick way for presenting design ideas	SA =1	SD =4	1.89	1.87
p2d	The computer was good to create my portfolio product	SA =1	SD =4	1.89	1.68
p2e	The computer was good for reflecting in the process doc	SA =1	SD =4	2.11	1.94
p2f	The computer was good for showing my skills in the portfolio	SA =1	SD =4	1.56	1.76
p2g	It was easy to follow the steps to create the portfolio	SA =1	SD =4	1.78	1.97
p2h	The steps of the portfolio helped me to develop my ideas	SA =1	SD =4	1.78	2.00
p2i	Overall, the computer is a good for portfolios	SA =1	SD =4	1.44	1.64
p2j	Overall, I was able to show what I can do in the portfolio.	SA =1	SD =4	1.78	1.86
p2k	It was better doing the portfolio on a computer than on paper	SA =1	SD =4	1.44	1.58
q5_com	Which of the following do you use at home? Computer	No=0	Yes=1	.89	.89
q5_dig	Digital camera	No=0	Yes=1	.89	.85
q5_vid	Vide0 camera	No=0	Yes=1	.89	.63
q5_mp3	MP3 player	No=0	Yes=1	.78	.95
q5_lap	Laptop	No=0	Yes=1	.78	.66
q5_gam	Game console	No=0	Yes=1	.89	.82
q5_mob	Mobile phone	No=0	Yes=1	.89	.93
q6	Do you have Internet access at home?	Dial up=2	Bdband=3	2.75	2.92
q7	How often do you use a computer at home?	Daily=1	Weekly=3	1.22	1.14
q8mon	Time (mins) spent on computers at school last week: Monday	0	120	107.86	97.57
q8tue	Tuesday	0	120	135.71	94.92
q8wed	Wednesday	0	180	97.86	98.64
q8thu	Thursday	0	120	92.86	85.36
q8fri	Friday	0	180	135.00	103.21
q9	Touch type?	Yes=1	No=2	1.00	1.27
q10a	Keep list of telephone numbers/addresses	I do=1	No=3	2.00	2.28
q10b	Draw a diagram or picture.	I do=1	No=3	1.75	2.01
q10c	Type an assignment for school	I do=1	No=3	1.00	1.10
q10d	Do a line graph or pie-diagram as part of an assignment.	I do=1	No=3	1.25	1.42
q10e	Send a letter to every club member or friend.	I do=1	No=3	1.13	1.54
q11a	Using computers makes the work at school more difficult	Yes=1	No=3	2.44	2.70
q11b	I enjoy using computers at school.	Yes=1	No=3	.89	1.30
q11c	I like to use a computer at home to do school work	Yes=1	No=3	1.00	1.26
q11d	I like to find things out for myself instead of being told.	Yes=1	No=3	1.33	1.75
q11e	Computers are good for the world	Yes=1	No=3	1.44	1.43
q12a	I feel confident working with computers	Yes=1	No=3	1.00	1.13
q12b	I'm good at using computers	Yes=1	No=3	1.25	1.30
q12c	I feel OK about trying a new problem on the computer	Yes=1	No=3	1.13	1.30
q12d	I usually do well with computers.	Yes=1	No=3	1.25	1.23
q12e	I could learn to program a computer.	Yes=1	No=3	1.75	1.71
q12f	Using a computer is very hard for me	Yes=1	No=3	3.00	2.90
q13_wp	Word processing: low ability =1 high ability =4	1	4	3.63	3.60
q13_ss	Spreadsheet: low ability =1 high ability =4	1	4	3.25	2.98
q13_db	Database: low ability =1 high ability =4	1	4	3.50	2.65
q13_sl	Slideshow: low ability =1 high ability =4	1	4	3.63	3.53
q13_em	Email: low ability =1 high ability =4	1	4	3.88	3.63

q13_fm	File Management: low ability =1 high ability =4	1	4	3.63	3.39
q13_in	Internet Research: low ability =1 high ability =4	1	4	4.00	3.73
q13_wa	Web Authoring: low ability =1 high ability =4	1	4	3.50	3.20
q13_dp	Digital Photography: low ability =1 high ability =4	1	4	3.88	3.49
q13_ie	Image Editing: low ability =1 high ability =4	1	4	3.50	3.54
q13_dv	Digital Video: low ability =1 high ability =4	1	4	3.50	2.90

MA Descriptive Statistics of Students' Questionnaire Responses (N=12)

(N=10)	**Abbreviated Questions**	**Min**	**Max**	**Mean**	**Pop. Mean**
Year	Student's schools year	11	11	11.00	11.13
Gender		Male=1	Female=2	1.00	1.20
q1a	How often have you done an exam on a computer before?	Often =1	None =4	3.40	3.20
q1b	How much more time would you need to get used to it?	Lots =1	None =4	2.70	2.80
q2a	It was easy to use the computer for doing the exam.	SA=1	SD=4	1.60	1.83
q2b	It was easy to use the comp. in the exam to develop ideas.	SA =1	SD =4	1.80	2.14
q2c	The computer was a quick way for presenting ideas in exam	SA =1	SD =4	1.50	1.96
q2d	The computer was good to create my logo, graphs, brochure.	SA =1	SD =4	1.40	1.41
q2e	The computer was good for reflecting on my design ideas	SA =1	SD =4	1.90	1.85
q2f	The computer was good for completing Sect A in the exam.	SA =1	SD =4	1.50	1.65
q2g	It was easy to follow the steps of the exam on the computer.	SA =1	SD =4	1.40	1.78
q2h	The steps of the exam helped me to develop my ideas.	SA =1	SD =4	1.90	1.92
q2i	The computer is a good tool for designing products exams.	SA =1	SD =4	1.50	1.73
q2j	Overall, I was able to show what I can do in the exam.	SA =1	SD =4	1.60	1.90
q2k	It was better doing the exam on a computer than on paper.	SA =1	SD =4	1.67	1.68
p1a	How often have you done a portfolio on a computer before?	Lots=1	None=4	2.70	2.47
p1b	How much more time would you need to get used to it?	Lots=1	None=4	2.30	2.67
p2a	It was easy to use the computer for doing the portfolio	SA =1	SD =4	1.70	1.89
p2b	It was easy to use the computer for my portfolio for ideas	SA =1	SD =4	1.80	2.05
p2c	The computer was a quick way for presenting design ideas	SA =1	SD =4	1.90	1.87
p2d	The computer was good to create my portfolio product	SA =1	SD =4	1.50	1.68
p2e	The computer was good for reflecting in the process doc	SA =1	SD =4	1.90	1.94
p2f	The computer was good for showing my skills in the portfolio	SA =1	SD =4	1.50	1.76
p2g	It was easy to follow the steps to create the portfolio	SA =1	SD =4	1.60	1.97
p2h	The steps of the portfolio helped me to develop my ideas	SA =1	SD =4	2.20	2.00
p2i	Overall, the computer is a good for portfolios	SA =1	SD =4	1.30	1.64
p2j	Overall, I was able to show what I can do in the portfolio.	SA =1	SD =4	1.50	1.86
p2k	It was better doing the portfolio on a computer than on paper	SA =1	SD =4	1.40	1.58
q5_com	Which of the following do you use at home? Computer	No=0	Yes=1	.70	.89
q5_dig	Digital camera	No=0	Yes=1	.60	.85

q5_vid	Vide0 camera	No=0	Yes=1	.40	.63
q5_mp3	MP3 player	No=0	Yes=1	.90	.95
q5_lap	Laptop	No=0	Yes=1	.60	.66
q5_gam	Game console	No=0	Yes=1	.70	.82
q5_mob	Mobile phone	No=0	Yes=1	.90	.93
q6	Do you have Internet access at home?	Dial up=2	Bdband=3	2.89	2.92
q7	How often do you use a computer at home?	Daily=1	Weekly=3	1.00	1.14
q8mon	Time (mins) spent on computers at school last week: Monday	0	120	96.67	97.57
q8tue	Tuesday	0	120	101.67	94.92
q8wed	Wednesday	0	180	100.56	98.64
q8thu	Thursday	0	120	105.33	85.36
q8fri	Friday	0	180	92.33	103.21
q9	Touch type?	Yes=1	No=2	1.00	1.27
q10a	Keep list of telephone numbers/addresses	I do=1	No=3	2.33	2.28
q10b	Draw a diagram or picture.	I do=1	No=3	2.11	2.01
q10c	Type an assignment for school	I do=1	No=3	1.11	1.10
q10d	Do a line graph or pie-diagram as part of an assignment.	I do=1	No=3	1.78	1.42
q10e	Send a letter to every club member or friend.	I do=1	No=3	1.44	1.54
q11a	Using computers makes the work at school more difficult	Yes=1	No=3	2.40	2.70
q11b	I enjoy using computers at school.	Yes=1	No=3	1.10	1.30
q11c	I like to use a computer at home to do school work	Yes=1	No=3	1.10	1.26
q11d	I like to find things out for myself instead of being told.	Yes=1	No=3	1.70	1.75
q11e	Computers are good for the world	Yes=1	No=3	1.00	1.43
q12a	I feel confident working with computers	Yes=1	No=3	1.11	1.13
q12b	I'm good at using computers	Yes=1	No=3	1.11	1.30
q12c	I feel OK about trying a new problem on the computer	Yes=1	No=3	1.22	1.30
q12d	I usually do well with computers.	Yes=1	No=3	1.22	1.23
q12e	I could learn to program a computer.	Yes=1	No=3	2.00	1.71
q12f	Using a computer is very hard for me	Yes=1	No=3	2.78	2.90
q13_wp	Word processing: low ability =1 high ability =4	1	4	3.44	3.60
q13_ss	Spreadsheet: low ability =1 high ability =4	1	4	2.67	2.98
q13_db	Database: low ability =1 high ability =4	1	4	2.33	2.65
q13_sl	Slideshow: low ability =1 high ability =4	1	4	3.89	3.53
q13_em	Email: low ability =1 high ability =4	1	4	3.78	3.63
q13_fm	File Management: low ability =1 high ability =4	1	4	3.22	3.39
q13_in	Internet Research: low ability =1 high ability =4	1	4	3.89	3.73
q13_wa	Web Authoring: low ability =1 high ability =4	1	4	3.56	3.20
q13_dp	Digital Photography: low ability =1 high ability =4	1	4	3.89	3.49
q13_ie	Image Editing: low ability =1 high ability =4	1	4	3.56	3.54
q13_dv	Digital Video: low ability =1 high ability =4	1	4	3.11	2.90

RA Descriptive Statistics of Students' Questionnaire Responses (N=14)

(N=14)	**Abbreviated Questions**	**Min**	**Max**	**Mean**	**Pop. Mean**
Year	Student's schools year	11	11	12.00	11.13
Gender		Male=1	Female=2	1.29	1.20
q1a	How often have you done an exam on a computer before?	Often =1	None =4	3.64	3.20

q1b	How much more time would you need to get used to it?	Lots =1	None =4	2.71	2.80
q2a	It was easy to use the computer for doing the exam.	SA=1	SD=4	1.79	1.83
q2b	It was easy to use the comp. in the exam to develop ideas.	SA =1	SD =4	1.86	2.14
q2c	The computer was a quick way for presenting ideas in exam	SA =1	SD =4	1.93	1.96
q2d	The computer was good to create my logo, graphs, brochure.	SA =1	SD =4	1.43	1.41
q2e	The computer was good for reflecting on my design ideas	SA =1	SD =4	1.86	1.85
q2f	The computer was good for completing Sect A in the exam.	SA =1	SD =4	1.64	1.65
q2g	It was easy to follow the steps of the exam on the computer.	SA =1	SD =4	1.79	1.78
q2h	The steps of the exam helped me to develop my ideas.	SA =1	SD =4	2.29	1.92
q2i	The computer is a good tool for designing products exams.	SA =1	SD =4	1.79	1.73
q2j	Overall, I was able to show what I can do in the exam.	SA =1	SD =4	1.79	1.90
q2k	It was better doing the exam on a computer than on paper.	SA =1	SD =4	1.64	1.68
p1a	How often have you done a portfolio on a computer before?	Lots=1	None=4	2.29	2.47
p1b	How much more time would you need to get used to it?	Lots=1	None=4	2.71	2.67
p2a	It was easy to use the computer for doing the portfolio	SA =1	SD =4	1.71	1.89
p2b	It was easy to use the computer for my portfolio for ideas	SA =1	SD =4	1.86	2.05
p2c	The computer was a quick way for presenting design ideas	SA =1	SD =4	1.86	1.87
p2d	The computer was good to create my portfolio product	SA =1	SD =4	1.79	1.68
p2e	The computer was good for reflecting in the process doc	SA =1	SD =4	1.86	1.94
p2f	The computer was good for showing my skills in the portfolio	SA =1	SD =4	1.79	1.76
p2g	It was easy to follow the steps to create the portfolio	SA =1	SD =4	1.64	1.97
p2h	The steps of the portfolio helped me to develop my ideas	SA =1	SD =4	1.86	2.00
p2i	Overall, the computer is a good for portfolios	SA =1	SD =4	1.71	1.64
p2j	Overall, I was able to show what I can do in the portfolio.	SA =1	SD =4	1.79	1.86
p2k	It was better doing the portfolio on a computer than on paper	SA =1	SD =4	1.57	1.58
q5_com	Which of the following do you use at home? Computer	No=0	Yes=1	.93	.89
q5_dig	Digital camera	No=0	Yes=1	.86	.85
q5_vid	Vide0 camera	No=0	Yes=1	.93	.63
q5_mp3	MP3 player	No=0	Yes=1	.71	.95
q5_lap	Laptop	No=0	Yes=1	.43	.66
q5_gam	Game console	No=0	Yes=1	.93	.82
q5_mob	Mobile phone	No=0	Yes=1	.93	.93
q6	Do you have Internet access at home?	Dial up=2	Bdband=3	2.77	2.92
q7	How often do you use a computer at home?	Daily=1	Weekly=3	1.38	1.14
q8mon	Time (mins) spent on computers at school last week: Monday	0	120	75.00	97.57
q8tue	Tuesday	0	120	30.00	94.92
q8wed	Wednesday	0	180	81.92	98.64
q8thu	Thursday	0	120	72.31	85.36
q8fri	Friday	0	180	78.46	103.21
q9	Touch type?	Yes=1	No=2	1.58	1.27
q10a	Keep list of telephone numbers/addresses	I do=1	No=3	2.85	2.28
q10b	Draw a diagram or picture.	I do=1	No=3	1.92	2.01
q10c	Type an assignment for school	I do=1	No=3	1.08	1.10
q10d	Do a line graph or pie-diagram as part of an assignment.	I do=1	No=3	1.54	1.42
q10e	Send a letter to every club member or friend.	I do=1	No=3	2.00	1.54
q11a	Using computers makes the work at school more difficult	Yes=1	No=3	2.62	2.70
q11b	I enjoy using computers at school.	Yes=1	No=3	1.69	1.30
q11c	I like to use a computer at home to do school work	Yes=1	No=3	1.23	1.26
q11d	I like to find things out for myself instead of being told.	Yes=1	No=3	1.77	1.75

q11e	Computers are good for the world	Yes=1	No=3	1.62	1.43
q12a	I feel confident working with computers	Yes=1	No=3	1.23	1.13
q12b	I'm good at using computers	Yes=1	No=3	1.23	1.30
q12c	I feel OK about trying a new problem on the computer	Yes=1	No=3	1.38	1.30
q12d	I usually do well with computers.	Yes=1	No=3	1.15	1.23
q12e	I could learn to program a computer.	Yes=1	No=3	1.85	1.71
q12f	Using a computer is very hard for me	Yes=1	No=3	2.85	2.90
q13_wp	Word processing: low ability =1 high ability =4	1	4	3.85	3.60
q13_ss	Spreadsheet: low ability =1 high ability =4	1	4	3.08	2.98
q13_db	Database: low ability =1 high ability =4	1	4	2.69	2.65
q13_sl	Slideshow: low ability =1 high ability =4	1	4	3.08	3.53
q13_em	Email: low ability =1 high ability =4	1	4	3.38	3.63
q13_fm	File Management: low ability =1 high ability =4	1	4	3.54	3.39
q13_in	Internet Research: low ability =1 high ability =4	1	4	3.77	3.73
q13_wa	Web Authoring: low ability =1 high ability =4	1	4	2.92	3.20
q13_dp	Digital Photography: low ability =1 high ability =4	1	4	3.38	3.49
q13_ie	Image Editing: low ability =1 high ability =4	1	4	3.38	3.54
q13_dv	Digital Video: low ability =1 high ability =4	1	4	3.08	2.90

WA Descriptive Statistics Of Students' Questionnaire Responses (N=14)

	Abbreviated Questions	Min	Max	Mean	Pop. Mean
Year	Student's schools year	11	11	11.00	11.13
Gender		Male=1	Female=2	1.00	1.20
q1a	How often have you done an exam on a computer before?	Often =1	None =4	3.21	3.20
q1b	How much more time would you need to get used to it?	Lots =1	None =4	2.57	2.80
q2a	It was easy to use the computer for doing the exam.	SA=1	SD=4	1.71	1.83
q2b	It was easy to use the comp. in the exam to develop ideas.	SA =1	SD =4	2.29	2.14
q2c	The computer was a quick way for presenting ideas in exam	SA =1	SD =4	1.93	1.96
q2d	The computer was good to create my logo, graphs, brochure.	SA =1	SD =4	1.43	1.41
q2e	The computer was good for reflecting on my design ideas	SA =1	SD =4	1.64	1.85
q2f	The computer was good for completing Sect A in the exam.	SA =1	SD =4	1.64	1.65
q2g	It was easy to follow the steps of the exam on the computer.	SA =1	SD =4	1.86	1.78
q2h	The steps of the exam helped me to develop my ideas.	SA =1	SD =4	1.64	1.92
q2i	The computer is a good tool for designing products exams.	SA =1	SD =4	1.79	1.73
q2j	Overall, I was able to show what I can do in the exam.	SA =1	SD =4	1.86	1.90
q2k	It was better doing the exam on a computer than on paper.	SA =1	SD =4	1.29	1.68
p1a	How often have you done a portfolio on a computer before?	Lots=1	None=4	2.21	2.47
p1b	How much more time would you need to get used to it?	Lots=1	None=4	2.36	2.67
p2a	It was easy to use the computer for doing the portfolio	SA =1	SD =4	1.71	1.89
p2b	It was easy to use the computer for my portfolio for ideas	SA =1	SD =4	1.86	2.05
p2c	The computer was a quick way for presenting design ideas	SA =1	SD =4	1.64	1.87
p2d	The computer was good to create my portfolio product	SA =1	SD =4	1.57	1.68
p2e	The computer was good for reflecting in the process doc	SA =1	SD =4	1.71	1.94
p2f	The computer was good for showing my skills in the portfolio	SA =1	SD =4	1.64	1.76
p2g	It was easy to follow the steps to create the portfolio	SA =1	SD =4	1.86	1.97
p2h	The steps of the portfolio helped me to develop my ideas	SA =1	SD =4	1.93	2.00
p2i	Overall, the computer is a good for portfolios	SA =1	SD =4	1.43	1.64
p2j	Overall, I was able to show what I can do in the portfolio.	SA =1	SD =4	1.57	1.86

p2k	It was better doing the portfolio on a computer than on paper	SA =1	SD =4	1.29	1.58
q5_com	Which of the following do you use at home? Computer	No=0	Yes=1	.93	.89
q5_dig	Digital camera	No=0	Yes=1	.79	.85
q5_vid	Vide0 camera	No=0	Yes=1	.57	.63
q5_mp3	MP3 player	No=0	Yes=1	.93	.95
q5_lap	Laptop	No=0	Yes=1	.71	.66
q5_gam	Game console	No=0	Yes=1	.79	.82
q5_mob	Mobile phone	No=0	Yes=1	.86	.93
q6	Do you have Internet access at home?	Dial up=2	Bdband=3	3.00	2.92
q7	How often do you use a computer at home?	Daily=1	Weekly=3	1.14	1.14
q8mon	Time (mins) spent on computers at school last week: Monday	0	120	107.14	97.57
q8tue	Tuesday	0	120	109.29	94.92
q8wed	Wednesday	0	180	97.86	98.64
q8thu	Thursday	0	120	93.21	85.36
q8fri	Friday	0	180	94.29	103.21
q9	Touch type?	Yes=1	No=2	1.29	1.27
q10a	Keep list of telephone numbers/addresses	I do=1	No=3	1.93	2.28
q10b	Draw a diagram or picture.	I do=1	No=3	2.07	2.01
q10c	Type an assignment for school	I do=1	No=3	1.07	1.10
q10d	Do a line graph or pie-diagram as part of an assignment.	I do=1	No=3	1.21	1.42
q10e	Send a letter to every club member or friend.	I do=1	No=3	1.50	1.54
q11a	Using computers makes the work at school more difficult	Yes=1	No=3	2.71	2.70
q11b	I enjoy using computers at school.	Yes=1	No=3	1.29	1.30
q11c	I like to use a computer at home to do school work	Yes=1	No=3	1.21	1.26
q11d	I like to find things out for myself instead of being told.	Yes=1	No=3	1.86	1.75
q11e	Computers are good for the world	Yes=1	No=3	1.07	1.43
q12a	I feel confident working with computers	Yes=1	No=3	1.14	1.13
q12b	I'm good at using computers	Yes=1	No=3	1.50	1.30
q12c	I feel OK about trying a new problem on the computer	Yes=1	No=3	1.29	1.30
q12d	I usually do well with computers.	Yes=1	No=3	1.21	1.23
q12e	I could learn to program a computer.	Yes=1	No=3	1.71	1.71
q12f	Using a computer is very hard for me	Yes=1	No=3	2.93	2.90
q13_wp	Word processing: low ability =1 high ability =4	1	4	3.71	3.60
q13_ss	Spreadsheet: low ability =1 high ability =4	1	4	3.50	2.98
q13_db	Database: low ability =1 high ability =4	1	4	2.71	2.65
q13_sl	Slideshow: low ability =1 high ability =4	1	4	3.21	3.53
q13_em	Email: low ability =1 high ability =4	1	4	3.43	3.63
q13_fm	File Management: low ability =1 high ability =4	1	4	3.43	3.39
q13_in	Internet Research: low ability =1 high ability =4	1	4	3.43	3.73
q13_wa	Web Authoring: low ability =1 high ability =4	1	4	2.79	3.20
q13_dp	Digital Photography: low ability =1 high ability =4	1	4	3.36	3.49
q13_ie	Image Editing: low ability =1 high ability =4	1	4	3.64	3.54
q13_dv	Digital Video: low ability =1 high ability =4	1	4	2.71	2.90

XA Descriptive Statistics Of Students' Questionnaire Responses (N=29).

	Abbreviated Questions	Min	Max	Mean	Pop. Mean
Year	Student's schools year	11	11	11.00	11.13
Gender		Male=1	Female=2	1.37	1.20

q1a	How often have you done an exam on a computer before?	Often =1	None =4	2.60	3.20
q1b	How much more time would you need to get used to it?	Lots =1	None =4	3.04	2.80
q2a	It was easy to use the computer for doing the exam.	SA=1	SD=4	1.72	1.83
q2b	It was easy to use the comp. in the exam to develop ideas.	SA =1	SD =4	2.12	2.14
q2c	The computer was a quick way for presenting ideas in exam	SA =1	SD =4	1.92	1.96
q2d	The computer was good to create my logo, graphs, brochure.	SA =1	SD =4	1.28	1.41
q2e	The computer was good for reflecting on my design ideas	SA =1	SD =4	1.72	1.85
q2f	The computer was good for completing Sect A in the exam.	SA =1	SD =4	1.68	1.65
q2g	It was easy to follow the steps of the exam on the computer.	SA =1	SD =4	1.84	1.78
q2h	The steps of the exam helped me to develop my ideas.	SA =1	SD =4	1.84	1.92
q2i	The computer is a good tool for designing products exams.	SA =1	SD =4	1.68	1.73
q2j	Overall, I was able to show what I can do in the exam.	SA =1	SD =4	1.96	1.90
q2k	It was better doing the exam on a computer than on paper.	SA =1	SD =4	1.60	1.68
p1a	How often have you done a portfolio on a computer before?	Lots=1	None=4	2.20	2.47
p1b	How much more time would you need to get used to it?	Lots=1	None=4	2.80	2.67
p2a	It was easy to use the computer for doing the portfolio	SA =1	SD =4	1.92	1.89
p2b	It was easy to use the computer for my portfolio for ideas	SA =1	SD =4	2.12	2.05
p2c	The computer was a quick way for presenting design ideas	SA =1	SD =4	1.72	1.87
p2d	The computer was good to create my portfolio product	SA =1	SD =4	1.60	1.68
p2e	The computer was good for reflecting in the process doc	SA =1	SD =4	1.80	1.94
p2f	The computer was good for showing my skills in the portfolio	SA =1	SD =4	1.72	1.76
p2g	It was easy to follow the steps to create the portfolio	SA =1	SD =4	2.20	1.97
p2h	The steps of the portfolio helped me to develop my ideas	SA =1	SD =4	2.04	2.00
p2i	Overall, the computer is a good for portfolios	SA =1	SD =4	1.56	1.64
p2j	Overall, I was able to show what I can do in the portfolio.	SA =1	SD =4	1.84	1.86
p2k	It was better doing the portfolio on a computer than on paper	SA =1	SD =4	1.44	1.58
q5_com	Which of the following do you use at home? Computer	No=0	Yes=1	.80	.89
q5_dig	Digital camera	No=0	Yes=1	.80	.85
q5_vid	Vide0 camera	No=0	Yes=1	.48	.63
q5_mp3	MP3 player	No=0	Yes=1	.96	.95
q5_lap	Laptop	No=0	Yes=1	.60	.66
q5_gam	Game console	No=0	Yes=1	.76	.82
q5_mob	Mobile phone	No=0	Yes=1	.96	.93
q6	Do you have Internet access at home?	Dial up=2	Bdband=3	3.00	2.92
q7	How often do you use a computer at home?	Daily=1	Weekly=3	1.00	1.14
q8mon	Time in mins spent using computers at school last week on: Monday	0	120	136.25	97.57
q8tue	Tuesday	0	120	121.25	94.92
q8wed	Wednesday	0	180	138.96	98.64
q8thu	Thursday	0	120	135.62	85.36
q8fri	Friday	0	180	134.79	103.21
q9	Touch type?	Yes=1	No=2	1.16	1.27
q10a	Keep list of telephone numbers/addresses	I do=1	No=3	2.43	2.28
q10b	Draw a diagram or picture.	I do=1	No=3	2.09	2.01
q10c	Type an assignment for school	I do=1	No=3	1.09	1.10
q10d	Do a line graph or pie-diagram as part of an assignment.	I do=1	No=3	1.39	1.42
q10e	Send a letter to every club member or friend.	I do=1	No=3	1.39	1.54
q11a	Using computers makes the work at school more difficult	Yes=1	No=3	2.70	2.70
q11b	I enjoy using computers at school.	Yes=1	No=3	1.39	1.30

q11c	I like to use a computer at home to do school work	Yes=1	No=3	1.22	1.26
q11d	I like to find things out for myself instead of being told.	Yes=1	No=3	1.83	1.75
q11e	Computers are good for the world	Yes=1	No=3	1.57	1.43
q12a	I feel confident working with computers	Yes=1	No=3	1.13	1.13
q12b	I'm good at using computers	Yes=1	No=3	1.09	1.30
q12c	I feel OK about trying a new problem on the computer	Yes=1	No=3	1.30	1.30
q12d	I usually do well with computers.	Yes=1	No=3	1.22	1.23
q12e	I could learn to program a computer.	Yes=1	No=3	1.65	1.71
q12f	Using a computer is very hard for me	Yes=1	No=3	2.91	2.90
q13_wp	Word processing: low ability =1 high ability =4	1	4	3.57	3.60
q13_ss	Spreadsheet: low ability =1 high ability =4	1	4	2.96	2.98
q13_db	Database: low ability =1 high ability =4	1	4	2.57	2.65
q13_sl	Slideshow: low ability =1 high ability =4	1	4	3.87	3.53
q13_em	Email: low ability =1 high ability =4	1	4	3.43	3.63
q13_fm	File Management: low ability =1 high ability =4	1	4	3.43	3.39
q13_in	Internet Research: low ability =1 high ability =4	1	4	3.43	3.73
q13_wa	Web Authoring: low ability =1 high ability =4	1	4	2.79	3.20
q13_dp	Digital Photography: low ability =1 high ability =4	1	4	3.36	3.49
q13_ie	Image Editing: low ability =1 high ability =4	1	4	3.64	3.54
q13_dv	Digital Video: low ability =1 high ability =4	1	4	2.71	2.90

ZA Descriptive Statistics of Students' Questionnaire Responses (N=16)

	Abbreviated Questions	**Min**	**Max**	**Mean**	**Pop. Mean**
Year	Student's schools year	11	11	11.00	11.13
Gender		Male=1	Female=2	1.00	1.20
q1a	How often have you done an exam on a computer before?	Often =1	None =4	3.06	3.20
q1b	How much more time would you need to get used to it?	Lots =1	None =4	2.63	2.80
q2a	It was easy to use the computer for doing the exam.	SA=1	SD=4	1.88	1.83
q2b	It was easy to use the comp. in the exam to develop ideas.	SA =1	SD =4	1.94	2.14
q2c	The computer was a quick way for presenting ideas in exam	SA =1	SD =4	2.06	1.96
q2d	The computer was good to create my logo, graphs, brochure.	SA =1	SD =4	1.25	1.41
q2e	The computer was good for reflecting on my design ideas	SA =1	SD =4	1.75	1.85
q2f	The computer was good for completing Sect A in the exam.	SA =1	SD =4	1.50	1.65
q2g	It was easy to follow the steps of the exam on the computer.	SA =1	SD =4	1.56	1.78
q2h	The steps of the exam helped me to develop my ideas.	SA =1	SD =4	1.56	1.92
q2i	The computer is a good tool for designing products exams.	SA =1	SD =4	1.44	1.73
q2j	Overall, I was able to show what I can do in the exam.	SA =1	SD =4	1.94	1.90
q2k	It was better doing the exam on a computer than on paper.	SA =1	SD =4	1.56	1.68
p1a	How often have you done a portfolio on a computer before?	Lots=1	None=4	2.00	2.47
p1b	How much more time would you need to get used to it?	Lots=1	None=4	2.69	2.67
p2a	It was easy to use the computer for doing the portfolio	SA =1	SD =4	1.69	1.89
p2b	It was easy to use the computer for my portfolio for ideas	SA =1	SD =4	1.81	2.05
p2c	The computer was a quick way for presenting design ideas	SA =1	SD =4	1.56	1.87
p2d	The computer was good to create my portfolio product	SA =1	SD =4	1.31	1.68
p2e	The computer was good for reflecting in the process doc	SA =1	SD =4	1.88	1.94
p2f	The computer was good for showing my skills in the portfolio	SA =1	SD =4	1.38	1.76
p2g	It was easy to follow the steps to create the portfolio	SA =1	SD =4	1.88	1.97
p2h	The steps of the portfolio helped me to develop my ideas	SA =1	SD =4	1.75	2.00

p2i	Overall, the computer is a good for portfolios	SA =1	SD =4	1.44	1.64
p2j	Overall, I was able to show what I can do in the portfolio.	SA =1	SD =4	1.63	1.86
p2k	It was better doing the portfolio on a computer than on paper	SA =1	SD =4	1.25	1.58
q5_com	Which of the following do you use at home? Computer	No=0	Yes=1	.94	.89
q5_dig	Digital camera	No=0	Yes=1	.88	.85
q5_vid	Vide0 camera	No=0	Yes=1	.50	.63
q5_mp3	MP3 player	No=0	Yes=1	1.00	.95
q5_lap	Laptop	No=0	Yes=1	.75	.66
q5_gam	Game console	No=0	Yes=1	.88	.82
q5_mob	Mobile phone	No=0	Yes=1	.88	.93
q6	Do you have Internet access at home?	Dial up=2	Bdband=3	2.88	2.92
q7	How often do you use a computer at home?	Daily=1	Weekly=3	1.00	1.14
q8mon	Time in mins spent using computers at school last week on: Monday	0	120	82.81	97.57
q8tue	Tuesday	0	120	80.31	94.92
q8wed	Wednesday	0	180	81.56	98.64
q8thu	Thursday	0	120	72.50	85.36
q8fri	Friday	0	180	97.50	103.21
q9	Touch type?	Yes=1	No=2	1.36	1.27
q10a	Keep list of telephone numbers/addresses	I do=1	No=3	1.81	2.28
q10b	Draw a diagram or picture.	I do=1	No=3	1.81	2.01
q10c	Type an assignment for school	I do=1	No=3	1.00	1.10
q10d	Do a line graph or pie-diagram as part of an assignment.	I do=1	No=3	1.19	1.42
q10e	Send a letter to every club member or friend.	I do=1	No=3	1.56	1.54
q11a	Using computers makes the work at school more difficult	Yes=1	No=3	2.94	2.70
q11b	I enjoy using computers at school.	Yes=1	No=3	1.13	1.30
q11c	I like to use a computer at home to do school work	Yes=1	No=3	1.38	1.26
q11d	I like to find things out for myself instead of being told.	Yes=1	No=3	1.44	1.75
q11e	Computers are good for the world	Yes=1	No=3	1.38	1.43
q12a	I feel confident working with computers	Yes=1	No=3	1.07	1.13
q12b	I'm good at using computers	Yes=1	No=3	1.13	1.30
q12c	I feel OK about trying a new problem on the computer	Yes=1	No=3	1.06	1.30
q12d	I usually do well with computers.	Yes=1	No=3	1.06	1.23
q12e	I could learn to program a computer.	Yes=1	No=3	1.56	1.71
q12f	Using a computer is very hard for me	Yes=1	No=3	2.94	2.90
q13_wp	Word processing: low ability =1 high ability =4	1	4	3.88	3.60
q13_ss	Spreadsheet: low ability =1 high ability =4	1	4	3.19	2.98
q13_db	Database: low ability =1 high ability =4	1	4	2.50	2.65
q13_sl	Slideshow: low ability =1 high ability =4	1	4	3.56	3.53
q13_em	Email: low ability =1 high ability =4	1	4	3.44	3.63
q13_fm	File Management: low ability =1 high ability =4	1	4	3.38	3.39
q13_in	Internet Research: low ability =1 high ability =4	1	4	3.69	3.73
q13_wa	Web Authoring: low ability =1 high ability =4	1	4	3.63	3.20
q13_dp	Digital Photography: low ability =1 high ability =4	1	4	3.50	3.49
q13_ie	Image Editing: low ability =1 high ability =4	1	4	3.88	3.54
q13_dv	Digital Video: low ability =1 high ability =4	1	4	2.63	2.90

Appendix L

AIT Task Assessment: Comparative-pairs Marking Criteria

Holistic Criterion: Brochure is effective for target customers through developed planning to incorporate all the required features and information, appropriate use of aesthetic effects on a theme, consistent and balanced layout, and professional look. [Evidenced across all components including evaluation]

Specific Criterion 1: Design Process. Product originates from planned design showing development of ideas and justification in reflection. [Focus on planning sheets at beginning of PDF, reflection at the end and the MP3 sound file]

Specific Criterion 2: Technical Proficiency. Demonstrable capability and facility with the range of required software (spreadsheet, logo, brochure). [Focus on features of graphs, logo and layout in brochure]

Specific Criterion 3: Design Principles. Creative application of appropriate design principles and elements such as alignment, balance, contrast, emphasis, harmony, proportion, proximity, repetition, unity, and white space. [Focus on brochure and logo].

Each of the 60 selected exemplars was allocated an identifying ‘Code’ (column 1). ‘Preferred’ (column 2) is the actual number of times the exemplar was preferred in all the ‘Involved’ (column 3) comparisons. The exemplars were ranked from best (most number of times preferred) to worst (least number of times preferred). ‘Estimate’ (column 4) is the exemplar location in logits (logarithmic units of measurement) and is the determinant of the rank order of the exemplars. ‘Std Err’ (column 5) is the standard error of measurement. ‘Outfit’ (column 6) is an index of whether the pattern was more or less Guttmann like, that is whether the exemplar was consistently rated. It is expected to have a value of about 1.00.

Appendix M

AIT Task Assessment: Comparative-pairs Marking Results by Criteria

Analysis 1: Ranking of student's work samples (marking criterion 1)

Code	Preferred	Involved	Estimate	Std Err	Outfit	Chi Sqr	Degrees Freedom
40	56	59	3.662	0.630	0.816	46.559	57.033
35	55	59	3.337	0.563	1.179	67.259	57.033
43	53	59	2.832	0.482	0.688	39.239	57.033
36	50	59	2.265	0.415	0.420	23.931	57.033
54	50	59	2.265	0.415	1.551	88.476	57.033
32	48	59	1.957	0.387	0.796	45.396	57.033
2	47	59	1.816	0.376	1.808	103.11	57.033
39	46	59	1.683	0.366	1.575	89.811	57.033
45	45	59	1.557	0.358	1.458	83.175	57.033
47	45	59	1.557	0.358	1.450	82.715	57.033
21	43	59	1.318	0.344	0.577	32.916	57.033
38	43	59	1.318	0.344	1.256	71.614	57.033
34	40	58	1.079	0.335	0.836	46.844	56.067
28	40	59	0.989	0.329	0.878	50.083	57.033
59	40	59	0.989	0.329	1.325	75.579	57.033
57	38	59	0.784	0.321	0.898	51.226	57.033
50	37	59	0.684	0.318	0.994	56.707	57.033
3	36	59	0.586	0.316	1.675	95.551	57.033
53	36	59	0.586	0.316	0.901	51.377	57.033
8	35	59	0.490	0.313	1.014	57.824	57.033
11	35	59	0.490	0.313	1.142	65.142	57.033
31	35	59	0.490	0.313	0.957	54.554	57.033
33	35	59	0.490	0.313	0.829	47.268	57.033
44	35	59	0.490	0.313	0.980	55.902	57.033
52	35	59	0.490	0.313	0.679	38.697	57.033
51	33	59	0.300	0.310	1.018	58.051	57.033
42	32	59	0.206	0.309	0.962	54.878	57.033
7	31	59	0.113	0.308	0.823	46.948	57.033
55	29	59	-0.072	0.308	0.824	46.985	57.033
29	28	59	-0.164	0.308	0.592	33.745	57.033
58	28	59	-0.164	0.308	0.723	41.219	57.033
48	27	59	-0.257	0.308	1.012	57.718	57.033
49	27	59	-0.257	0.308	0.977	55.73	57.033
10	26	59	-0.350	0.309	1.230	70.123	57.033
13	26	59	-0.350	0.309	0.804	45.843	57.033
27	26	59	-0.350	0.309	1.299	74.098	57.033
16	25	59	-0.443	0.310	1.075	61.297	57.033
4	24	59	-0.537	0.311	0.866	49.365	57.033
41	24	59	-0.537	0.311	1.404	80.057	57.033
60	23	58	-0.618	0.315	1.165	65.299	56.067
6	23	59	-0.632	0.313	1.809	103.185	57.033
46	23	59	-0.632	0.313	0.626	35.693	57.033
1	22	59	-0.729	0.315	1.066	60.805	57.033
56	22	59	-0.729	0.315	0.792	45.153	57.033
24	21	59	-0.826	0.318	1.186	67.614	57.033

9	19	59	-1.027	0.324	1.180	67.323	57.033
30	19	59	-1.027	0.324	1.006	57.396	57.033
14	18	59	-1.131	0.328	0.794	45.263	57.033
5	16	59	-1.347	0.338	0.635	36.19	57.033
15	16	59	-1.347	0.338	1.192	67.997	57.033
20	16	59	-1.347	0.338	1.451	82.745	57.033
17	13	59	-1.699	0.359	1.116	63.63	57.033
37	13	59	-1.699	0.359	1.036	59.106	57.033
22	12	59	-1.826	0.368	2.650	151.108	57.033
25	12	59	-1.826	0.368	1.164	66.385	57.033
26	12	59	-1.826	0.368	0.978	55.752	57.033
18	8	59	-2.417	0.421	0.987	56.312	57.033
12	7	59	-2.595	0.442	0.747	42.582	57.033
19	6	59	-2.792	0.468	0.500	28.514	57.033
23	4	59	-3.271	0.547	0.837	47.749	57.033

Analysis 2: Analysis of the spread of the exemplars (marking criterion 1)

A Separation Index was calculated as an indicator as to whether or not the exemplars were sufficiently diverse in quality so as to assure a broad enough range for the purposes of comparison. It is given as a number from 0 to 1. Values closer to 1.00 are more desirable. If the value is close to 0.00 (up to about 0.3 or 0.4) the range is too narrow. If it is above about 0.7 the separation is reasonable and if it is above 0.8, the separation is good. In this case, the Separation Index for AIT examination practical component criterion 1 was 0.940 indicating a very good spread of quality in the exemplars.

Analysis 3: Inter-rater reliability analysis (marking criterion 1)

Intra-rater reliability analysis was undertaken in order to assess individual judge consistency in relation to the judgments of the other judges in the group. The “Outfit” statistic, in this instance, should be between 0.5 and 1.5

Judge	**Outfit**
1	1.057
2	0.800
3	1.259
4	1.145
5	1.007

The group reliability is defined as the average of the individual rater reliability indices. For Criterion 1 the group reliability was 1.05.

Analysis 1: Ranking of student's work samples (marking criterion 2)

Code	Preferred	Involved	Estimate	Std Err	Outfit	Chi Sqr	Degrees Freedom
40	55	59	3.492	0.562	1.861	106.119	57.033
45	54	59	3.222	0.517	2.382	135.86	57.033
59	54	59	3.222	0.517	0.421	24.002	57.033
54	53	59	2.987	0.484	0.399	22.765	57.033
53	50	59	2.410	0.421	0.638	36.409	57.033
33	47	59	1.946	0.384	0.565	32.225	57.033
36	47	59	1.946	0.384	1.188	67.754	57.033
44	46	59	1.807	0.375	0.793	45.218	57.033
57	45	59	1.674	0.367	0.570	32.536	57.033
58	45	59	1.674	0.367	1.107	63.143	57.033
2	44	59	1.546	0.360	1.345	76.702	57.033
31	44	59	1.546	0.360	1.770	100.957	57.033
35	44	59	1.546	0.360	0.679	38.704	57.033
60	42	58	1.413	0.355	0.688	38.563	56.067
39	42	59	1.302	0.349	0.931	53.106	57.033
43	42	59	1.302	0.349	1.719	98.048	57.033
52	42	59	1.302	0.349	0.442	25.219	57.033
55	42	59	1.302	0.349	0.899	51.255	57.033
21	40	59	1.072	0.340	1.292	73.689	57.033
11	36	59	0.641	0.327	0.670	38.232	57.033
32	35	59	0.538	0.325	0.657	37.45	57.033
50	35	59	0.538	0.325	0.873	49.763	57.033
3	33	59	0.335	0.321	1.035	59.043	57.033
28	32	59	0.235	0.320	0.838	47.766	57.033
47	32	59	0.235	0.320	1.352	77.091	57.033
46	31	59	0.136	0.318	1.670	95.243	57.033
14	30	59	0.037	0.318	1.244	70.963	57.033
29	30	59	0.037	0.318	0.570	32.529	57.033
27	29	59	-0.061	0.317	0.924	52.721	57.033
38	28	59	-0.159	0.317	0.936	53.407	57.033
1	27	59	-0.257	0.317	0.881	50.256	57.033
42	27	59	-0.257	0.317	1.256	71.633	57.033
15	26	59	-0.355	0.317	1.202	68.574	57.033
30	25	59	-0.454	0.318	1.490	85.004	57.033
51	25	59	-0.454	0.318	0.797	45.457	57.033
56	25	59	-0.454	0.318	0.830	47.325	57.033
23	24	59	-0.553	0.319	1.100	62.722	57.033
49	23	59	-0.653	0.321	0.583	33.262	57.033
4	22	59	-0.753	0.323	0.903	51.494	57.033
41	22	59	-0.753	0.323	1.200	68.438	57.033
10	21	59	-0.855	0.325	0.601	34.268	57.033
16	21	59	-0.855	0.325	0.999	56.949	57.033
20	21	59	-0.855	0.325	1.477	84.239	57.033
17	20	59	-0.959	0.328	0.815	46.493	57.033
34	19	58	-1.056	0.332	0.737	41.324	56.067
13	19	59	-1.065	0.331	0.568	32.402	57.033
24	19	59	-1.065	0.331	4.358	248.573	57.033
8	18	59	-1.173	0.335	0.610	34.813	57.033

9	17	59	-1.283	0.340	0.928	52.953	57.033
5	16	59	-1.397	0.345	0.732	41.763	57.033
6	16	59	-1.397	0.345	0.663	37.806	57.033
48	16	59	-1.397	0.345	0.600	34.195	57.033
37	15	59	-1.515	0.351	1.595	90.963	57.033
12	13	59	-1.764	0.367	0.795	45.344	57.033
25	13	59	-1.764	0.367	0.616	35.132	57.033
22	11	59	-2.039	0.389	2.103	119.947	57.033
7	8	59	-2.527	0.438	0.526	30	57.033
18	5	59	-3.181	0.532	0.458	26.131	57.033
26	5	59	-3.181	0.532	1.072	61.151	57.033
19	1	59	-4.947	1.033	0.339	19.33	57.033

Analysis 2: Analysis of the spread of the exemplars (marking criterion 2)

The Separation Index for AIT examination practical component criterion 2 was 0.946 indicating a very good spread of quality in the exemplars.

Analysis 3: Intra-rater reliability analysis (marking criterion 2)

Judge	Outfit
1	0.884
2	0.685
3	1.022
4	1.426
5	1.092

For Criterion 2 the group reliability was 1.02

Analysis 1: Ranking of student's work samples (marking criterion 3)

Code	Preferred	Involved	Estimate	Std Err	Outfit	Chi Sqr	Degrees Freedom
40	56	59	4.009	0.640	0.242	13.83	57.033
36	54	59	3.388	0.532	1.635	93.242	57.033
2	53	58	3.386	0.533	10.064	564.204	56.067
35	52	59	2.915	0.475	2.386	136.06	57.033
45	52	59	2.915	0.475	1.137	64.818	57.033
54	52	59	2.915	0.475	0.561	31.977	57.033
44	50	59	2.521	0.438	0.888	50.665	57.033
59	48	59	2.177	0.411	1.693	96.573	57.033
43	47	59	2.019	0.400	1.075	61.306	57.033
39	46	59	1.868	0.391	0.739	42.155	57.033
31	45	59	1.724	0.382	1.191	67.917	57.033
53	44	59	1.586	0.375	1.194	68.087	57.033
33	43	59	1.453	0.368	0.732	41.769	57.033
57	43	59	1.453	0.368	0.432	24.664	57.033
21	42	59	1.324	0.362	0.722	41.149	57.033
3	41	58	1.318	0.363	0.860	48.234	56.067

38	40	59	1.078	0.351	0.795	45.357	57.033
55	39	59	0.959	0.347	0.780	44.482	57.033
58	38	59	0.844	0.343	1.416	80.759	57.033
52	36	59	0.621	0.336	0.820	46.776	57.033
11	33	59	0.300	0.328	0.673	38.362	57.033
28	33	59	0.300	0.328	0.739	42.168	57.033
32	33	59	0.300	0.328	0.524	29.873	57.033
50	33	59	0.300	0.328	1.115	63.572	57.033
51	33	59	0.300	0.328	0.969	55.275	57.033
4	32	59	0.196	0.326	0.654	37.313	57.033
27	30	59	-0.010	0.323	0.845	48.18	57.033
47	30	59	-0.010	0.323	1.624	92.608	57.033
34	29	58	-0.042	0.326	1.673	93.815	56.067
29	27	59	-0.313	0.321	0.679	38.742	57.033
1	26	59	-0.413	0.321	0.915	52.164	57.033
48	26	59	-0.413	0.321	1.075	61.315	57.033
30	25	59	-0.514	0.322	2.189	124.832	57.033
46	25	59	-0.514	0.322	1.153	65.772	57.033
24	24	59	-0.615	0.323	1.491	85.038	57.033
9	23	59	-0.717	0.324	0.866	49.403	57.033
14	23	59	-0.717	0.324	0.931	53.083	57.033
15	23	59	-0.717	0.324	1.157	65.972	57.033
60	22	58	-0.787	0.329	2.351	131.818	56.067
16	22	59	-0.820	0.326	1.177	67.12	57.033
10	21	59	-0.924	0.328	0.813	46.344	57.033
23	21	59	-0.924	0.328	0.813	46.384	57.033
42	21	59	-0.924	0.328	0.750	42.796	57.033
8	20	59	-1.030	0.331	0.657	37.46	57.033
13	20	59	-1.030	0.331	0.573	32.682	57.033
17	20	59	-1.030	0.331	1.701	97.024	57.033
49	20	59	-1.030	0.331	0.644	36.755	57.033
56	19	59	-1.137	0.334	1.763	100.521	57.033
5	17	59	-1.360	0.343	0.694	39.601	57.033
7	17	59	-1.360	0.343	0.711	40.557	57.033
41	16	59	-1.475	0.348	1.135	64.715	57.033
6	15	59	-1.595	0.354	0.691	39.418	57.033
25	14	58	-1.713	0.362	0.508	28.482	56.067
22	14	58	-1.718	0.362	2.727	152.908	56.067
12	12	59	-1.984	0.380	0.770	43.916	57.033
37	8	59	-2.618	0.439	0.421	24.035	57.033
20	7	59	-2.812	0.462	9.915	565.443	57.033
18	6	59	-3.026	0.491	0.594	33.866	57.033
26	4	59	-3.550	0.576	0.578	32.954	57.033
19	2	59	-4.320	0.756	0.351	20.041	57.033

Analysis 2: Analysis of the spread of the exemplars (marking criterion 3)

The Separation Index for AIT examination practical component criterion 3 was 0.951 indicating a very good spread of quality in the exemplars.

Analysis 3: Intra-rater reliability analysis (marking criterion 3)

Judge	Outfit
1	1.342
2	0.831
3	0.903
4	2.517
5	0.972

For Criterion 3 the group reliability was 1.31

Analysis 1: Ranking of student's work samples (marking criterion holistic)

Code	Preferred	Involved	Estimate	Std Err	Outfit	Chi Sqr	Degrees Freedom
40	56	59	4.198	0.640	0.184	10.52	57.033
36	55	59	3.861	0.576	1.531	87.324	57.033
44	52	59	3.104	0.476	0.357	20.339	57.033
54	52	59	3.104	0.476	0.535	30.522	57.033
45	51	58	3.102	0.476	1.849	103.692	56.067
35	51	59	2.898	0.456	1.388	79.139	57.033
59	51	59	2.898	0.456	0.501	28.567	57.033
43	48	59	2.358	0.416	1.098	62.63	57.033
2	47	59	2.196	0.406	1.992	113.632	57.033
53	47	59	2.196	0.406	0.663	37.787	57.033
39	46	59	2.041	0.398	0.725	41.334	57.033
33	45	59	1.892	0.390	0.322	18.342	57.033
21	44	59	1.747	0.384	0.970	55.33	57.033
31	44	59	1.747	0.384	1.677	95.65	57.033
55	43	59	1.607	0.378	0.642	36.616	57.033
3	42	59	1.470	0.373	3.043	173.522	57.033
52	41	59	1.337	0.369	0.400	22.837	57.033
57	41	59	1.337	0.369	0.357	20.369	57.033
58	40	59	1.207	0.365	0.671	38.294	57.033
28	39	59	1.080	0.361	0.633	36.099	57.033
38	39	59	1.080	0.361	0.829	47.271	57.033
50	37	59	0.832	0.354	1.186	67.666	57.033
11	34	59	0.476	0.346	0.643	36.646	57.033
32	34	59	0.476	0.346	0.635	36.243	57.033
47	34	59	0.476	0.346	2.042	116.438	57.033
27	30	59	0.020	0.339	0.973	55.516	57.033
51	29	59	-0.092	0.338	0.775	44.209	57.033
1	28	59	-0.203	0.338	1.447	82.551	57.033
4	27	59	-0.314	0.337	0.873	49.764	57.033
46	27	59	-0.314	0.337	0.725	41.353	57.033
10	26	59	-0.424	0.337	0.926	52.816	57.033
30	26	59	-0.424	0.337	2.811	160.291	57.033
15	25	59	-0.535	0.338	1.481	84.476	57.033
29	25	59	-0.535	0.338	0.674	38.457	57.033

60	25	59	-0.535	0.338	1.331	75.92	57.033
34	24	59	-0.646	0.339	0.819	46.715	57.033
49	23	59	-0.758	0.340	0.527	30.052	57.033
16	22	59	-0.871	0.342	0.856	48.812	57.033
17	22	59	-0.871	0.342	0.844	48.146	57.033
42	22	59	-0.871	0.342	0.676	38.534	57.033
48	22	59	-0.871	0.342	0.927	52.855	57.033
56	21	59	-0.985	0.344	1.609	91.788	57.033
8	20	59	-1.100	0.347	0.591	33.695	57.033
9	20	59	-1.100	0.347	0.948	54.052	57.033
14	20	59	-1.100	0.347	1.203	68.615	57.033
24	19	59	-1.218	0.350	1.893	107.95	57.033
13	18	59	-1.339	0.354	0.539	30.768	57.033
23	18	59	-1.339	0.354	0.954	54.384	57.033
5	17	59	-1.462	0.359	0.778	44.357	57.033
41	16	58	-1.588	0.365	1.169	65.528	56.067
6	16	59	-1.589	0.365	1.080	61.577	57.033
25	16	59	-1.589	0.365	0.554	31.578	57.033
7	12	59	-2.149	0.399	0.539	30.739	57.033
12	10	59	-2.474	0.425	1.936	110.397	57.033
37	7	59	-3.055	0.482	2.057	117.292	57.033
18	6	59	-3.285	0.508	0.664	37.849	57.033
20	5	59	-3.543	0.542	0.601	34.282	57.033
26	5	59	-3.543	0.542	0.553	31.564	57.033
22	4	59	-3.837	0.586	0.830	47.321	57.033
19	3	59	-4.184	0.650	0.488	27.812	57.033

Analysis 2: Analysis of the spread of the exemplars (marking criterion holistic)

The Separation Index for AIT examination practical component holistic was 0.958 indicating a very good spread of quality in the exemplars.

Analysis 3: Intra-rater reliability analysis (Holistic)

Judge	Outfit
1	0.91
2	0.602
3	1.016
4	1.367
5	1.147

For Holistic the group reliability was 1.01