A heuristic framework for the determination of the critical elements in authentic assessment

Kevin Hugh Ashford-Rowe
University of Wollongong

A HEURISTIC FRAMEWORK FOR THE DETERMINATION OF THE CRITICAL ELEMENTS IN AUTHENTIC ASSESSMENT

A thesis submitted in partial fulfilment of the requirements for the award of the degree

DOCTOR OF EDUCATION

FROM

UNIVERSITY OF WOLLONGONG

BY

KEVIN HUGH ASHFORD-ROWE, BACHELOR OF ARTS (HONOURS), POST GRADUATE CERTIFICATE IN EDUCATION, GRADUATE DIPLOMA IN MULTIMEDIA, MASTER OF PROFESSIONAL STUDIES, MASTER OF EDUCATION

FACULTY OF EDUCATION

2009

DECLARATION

I, Kevin H. Ashford-Rowe, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Doctor of Education, in the Faculty of Education, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

Kevin H. Ashford-Rowe

23 January 2009.

TABLE OF CONTENTS

Declaration ..... **ii**

List of tables ..... **vii**

List of Figures ..... **vii**

Abstract ..... **viii**

Acknowledgments ..... **ix**

Chapter 1: Introduction **1**

Background to the study — Authenticity in educational assessment 1
Assessment, authenticity and educational technology 4
Research questions and the study 8
The organisation of the thesis 10

Chapter 2: Authentic Assessment: A General review of the literature **12**

Previous findings 12
Assessment 12
Assessment in higher education 17
What is authentic assessment 19
Assessment and educational technology 23
Characteristics of authentic assessment 31

Chapter 3: Research Design and Methodology **37**

Introduction 37
Design-based research 37
PHASE 1: Exploration of the problem 43
PHASE 2: Development of a solution 44
PHASE 3: Implementation and evaluation 50
PHASE 4: Presentation of findings 58
Summary of the research plan 59
Conclusion 62

Chapter 4: An Effective Model for Task Design in Flexible Learning Environments **63**

Analysis of the elements of authentic assessment 63
Practitioner feedback 66
Evolving and further developing the critical elements 69
1. Degree of challenge(s) presented to the assessed student 70
2. Performance, or product, as final assessment outcome 72
3. Transfer of learning (skills/knowledge/attitude) required 74
4. Critical reflection and self-assessment or evaluation required 75
5. Accuracy in product or performance, and fidelity of assessment environment, is displayed 76
6. Fidelity of assessment tools used 78

7. Discussion and feedback required 79
8. Collaboration required 80
Expert review 81
Expert Reviewer 1 82
Expert Reviewer 2 85
Expert Reviewer 3 88
Summary of feedback from expert reviewers 90
1. Degree of challenge(s) presented to the assessed student 91
2. Performance, or product, as final assessment outcome 92
3. Transfer of learning (skills/knowledge/attitude) required 92
4. Critical reflection and self-assessment or evaluation required 93
5. Accuracy in product or performance, and fidelity of assessment environment, is displayed 93
6. Fidelity of assessment tools used 94
7. Discussion and feedback required 94
8. Collaboration required 95
Revision of critical elements from expert review 95
From critical elements to critical questions — A summary 98
The critical questions 98

Chapter 5: Applying the Critical Questions of Authentic Assessment in the Design of a Learning Module 100

Development of — *Evaluating Educational Multimedia* 100
Introduction 100
The re-design of *Evaluating Educational Multimedia* 101
1. To what extent does the assessment activity challenge the assessed student? 105
2. Is a performance, or product, required as a final assessment outcome? 106
3. Does the assessment activity require that transfer of learning has occurred, by means of demonstration of skill? 107
4. Does the assessment activity require that metacognition is demonstrated, by means of critical reflection, self-assessment or evaluation? 108
5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder? 110
6. Is fidelity required in the assessment environment? And in the assessment tools (actual or simulated)? 114
7. Does the assessment activity require discussion and feedback? 115
8. Does the assessment activity require that students collaborate? 116
9. Description of how the critical questions were applied in the design and structure of the learning outcomes and assessment criteria of Module 10 117
The role of formative assessment in the redesign of the module 133
The application of the elements to the learning environment 135
Conclusion 136

Chapter 6: Learners' Responses to Authentic Assessment 139

Learning Module Implementation 140
Learning Module Evaluation and Analysis 140

Method of implementation 140
The method of analysis 141
Applying the constant comparative method 142
Analysis of responses 143
Researcher's observation on students' responses by data source 144
Interview 144
Observation 145
Video 146
Notes on students' performance made on observation during the delivery of the module 147
Notes on student's performance made on researcher review of the video content recorded during the delivery of the module 150
The student's response to the critical questions 152
1. To what extent does the assessment activity challenge the assessed student? 152
2. Is a performance, or product, required as a final assessment outcome? 155
3. Does the assessment activity require that transfer of learning has occurred, by means of demonstration of skill? 157
4. Does the assessment activity require that metacognition, is demonstrated, by means of critical reflection, self-assessment or evaluation? 159
5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder? 160
6. Is fidelity required in the assessment environment? And the assessment tools (actual or simulated)? 162
7. Does the assessment activity require discussion and feedback? 163
8. Does the assessment activity require that students collaborate? 165
Summary of the student's response to the application of the critical questions in the redesign of Module 10 167
The student's response to the assessment activity 170
Discussion 172

Chapter 7: Discussion 181

Research questions — Data analysis 183
1. To what extent does the assessment activity challenge the assessed student? 186
2. Is a performance or product required as a final assessment outcome? 188
3. Does the assessment activity require that transfer of learning has occurred, by means of demonstration of skill? 189
4. Does the assessment activity require that metacognition is demonstrated by means of critical reflection, self-assessment or evaluation? 190
5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder? 191
6. Is fidelity required in the assessment environment? And the assessment tools (actual or simulated)? 193
7. Does the assessment activity require discussion and feedback? 194
8. Does the assessment activity require that students collaborate? 195

Summary of student response and impact on the critical questions..... 199

Chapter 8: Conclusion 201

Introduction 201
Summary and review of process 201
PHASE 1: Exploration of the problem 204
PHASE 2: Development of a solution 204
PHASE 3: Implementation and evaluation 206
PHASE 4: Presentation of findings 207
Description of the principles 207
Findings of the study 209
Principal research question 210
Subordinate research question 1 211
Subordinate research question 2 215
Conclusion 217
Limitations of the study 218
Recommendations for further research 219

References 221

APPENDIX 1 Expert Reviewer Interview Questionnaire 239

APPENDIX 2 Student Evaluation Questionnaire 243

APPENDIX 3 Student interview Questionnaire 247

APPENDIX 4 Computer Based Learning Practitioners Course — Module 10 — Evaluating educational multimedia 252

LIST OF TABLES

Table 3.1: The way in which the stages of the design-based research process are applied in this study. **42**

Table 3.2: Summary of the research plan **61**

Table 4.1: Researcher's synthesis of the elements of authentic assessment from the literature **64**

Table 4.2 Researcher's translation of characteristics to critical elements of authentic assessment with practitioner feedback **67**

Table 4.3: Revision of critical elements from expert reviewer feedback to produce the critical questions **96**

Table 5.1: Proposed application of the critical questions to the re-design of Module 10 **104**

Table 6.1: Student Feedback on the Critical Elements **179**

Table 7.1: Consideration of the students'responses with reference to the research questions **183**

Table 7.2: Student Feedback on the Critical Questions **198**

Table 8.1: Stages of the design-based research process in this study **203**

LIST OF FIGURES

Figure 3.1: Design-based research (2006, p. 59) **39**

Figure 5.1: Apply the process of educational multimedia evaluation to the Army's Training Technology Centre developed Computer Based Learning Practitioners Course **110**

Figure 5.2: Apply the process of educational multimedia evaluation to a TTC developed CBLP — Trainees will construct their own **113**

Figure 5.3: Process of multimedia evaluation model assessment activity **117**

ABSTRACT

Higher Education is currently undergoing a period of significant challenge and transformation. It is likely that these challenges will, in a comparatively short period of time, lead to changes in the ways in which the higher education experience is both mediated and accessed. These changes have arisen as a result of a number of factors, including the information revolution, and the consequent pace of technological innovation, the increased demand from both employers and government for a more highly skilled workforce and the desire to increase and make more accessible the higher education experience to an increasing proportion of the overall population.

All of this has impacted upon the ways in which the higher education experience is represented, and in turn, by which students gain access to the knowledge and skills that will underpin their ability to both learn and perform. Higher education is increasingly being challenged to demonstrate its continued value to the broader community, especially employers, by ensuring that it provides capable, competent and informed citizens adequate to the challenges of a twenty-first century lifetime. If these principles are considered drivers for change, then it is important that the higher education sector can continue to demonstrate its ongoing value to the students who undertake it.

It is against this background that this study was developed with the purpose of identifying from the literature, and then to codify into an applicable framework, the critical elements that would determine an assessment as being authentic. The study took as its starting point the importance, in the current educational context, of being able to determine the elements that define an educational experience as being an authentic one. The research commenced with a review of the literature to identify and collate those elements that had been identified by previous researchers in the field. Next these elements, once refined iteratively in practice, were developed into a framework that could be applied by the designer of instruction and assessment, in order to ascertain whether such a framework could be used to support the design of a more authentic assessment experience. This framework was then applied in practice and the student's response to the learning and assessment designed according to these elements was evaluated, and the elements were further reviewed and revised upon the basis of this data. Thus the study was conducted in four phases, in the first of which the researcher explored the problem, in the second the researcher sought the development of a solution, and in the third phase this solution was implemented and evaluated, the findings were presented in the final phase.

The findings of this study suggest that not only is it possible to codify those elements critical to the determination of authenticity into such a framework, but moreover, it is possible to systematically apply them in the design of assessment activity. Thus the implication of this research for educators and educational designers who seek to meet a requirement for workplace relevance in the design of their education and assessment activities is that they will have a better opportunity to both identify and then apply specific design principles that will assist them in the better development of assessment outcomes with a clearer workplace applicability.

ACKNOWLEDGMENTS

In the completion of this work, I acknowledge the help, guidance and assistance freely given to me by a number of people over many years, from both within and outside of the field of education.

First, and foremost, I thank my wife, Tina, for her love and unstinting support, and my children, Holly, Jamie and Sam, for their continued interest and the questions that have ensured that I saw this through. I also thank all of them for the understanding that has enabled me to sacrifice time with them to the completion of this study.

I thank my parents Ken and Catherine whose love, support, guidance and belief ensured that both I, and my brothers, Jeremy, Alan and Ian, have been adequate to the joys and challenges that a life can present.

From an educational perspective, I acknowledge the inspiration and assistance given to me by many people, in particular, Robert Pepper and Lotte Deeble, Bosvigo County Primary, Cornwall (1971–1974), Rosemary King, Penweathers County Secondary and Richard Lander Comprehensive, Cornwall (1974–1979) and Margaret Garland and David Worley of Cornwall College, Cornwall (1979–1982).

I also sincerely acknowledge the assistance and advice of Professor John Hedberg and Professor Barry Harper.

Finally, I am deeply indebted to, and thank, my supervisors Associate Professor Jan Herrington and Doctor Christine Brown, who have provided me with both support and guidance over the last several years, and without whose vision and the means to express it, I would not have come close to completing this work.

I dedicate this work to my father, Leslie Kenneth Rowe, Kernowyon, (1931–1992) who understood the true value and importance of education as the enabler that can allow us to fulfil our potential and raise ourselves up.

My thanks to Jill Ryan for her excellent assistance in proof reading this thesis.

‘Then said a teacher, Speak to us of Teaching.

And he said:

No man can reveal to you aught but that which already lies half asleep in the dawning of your knowledge.

The teacher who walks in the shadow of the temple, among his followers, gives not of his wisdom but rather of his faith and his lovingness.

If he is indeed wise he does not bid you to enter the house of his wisdom, but rather leads you to the threshold of your own mind.’ (Khalil Gibran, 1923, p. 74)

CHAPTER 1: INTRODUCTION

Background to the study — Authenticity in educational assessment

An increasingly common theme throughout the later years of secondary education as well as across the tertiary education sector is that of professional preparedness and the development of generic workplace competencies that will fit a student for future employment. Nowhere is this more apparent than in the Australian higher education sector where the development and embedding of such generic skills, designed to ensure that a graduate is 'workplace ready', have become a focus for curriculum re-design.

One of the current trends evident in higher education is the 'increased interest in linking employment outcomes to higher education' (Chalmers, 2007, p. 7). That is not to say that Australian universities, particularly through programs and courses with a requirement to meet professional standards, have not always been aware that many students are studying for workplace readiness or advancement, as opposed to seeking careers as pure researchers. Rather, it is that in an increasingly competitive higher education market, many universities both pride and market themselves on the formal expression of their work/study links in the design and structure of the courses that they deliver. Indeed the Australian Learning and Teaching Council (ALTC) notes this requirement for curriculum renewal in its first objective to:

- Promote and support strategic change in higher education institutions for the enhancement of learning and teaching, including curriculum development and assessment (p. 2)

The role that assessment plays in the determination of professional, vocational and academic achievement is critical to this strategic change in the education process and its outcome, irrespective of the means of delivery. Allied to this change has been a shift from the application of a predominantly behavioural pedagogy to a constructivist learning paradigm. Such a paradigm gives a much greater emphasis to the value of situating the learning experience in an authentic context to enhance learning and teaching. In order to inform the broader field of higher education, situated as it currently is in an increasingly technically literate communication environment, this thesis seeks to harness principles of authenticity to guide the design and development of more meaningful assessment activity.

Over the past twenty five years education has undergone significant changes in the ways in which curriculum can be both facilitated and constructed. DeCastro-Ambrosetti and Cho's (2005, p.58), research indicates that students have become more culturally diverse and continue to be diverse in their learning needs and learning styles. This requirement to acknowledge the impact of changing curriculum, alongside the increased importance of recognising student diversity, has occurred over a period of time when the impact of external control over education delivery has also increased.

Whilst much of this change has arisen from the desire of governments to establish national frameworks for the delivery, assessment and accreditation of education and training, another factor impacting upon the extent of the change has been the prevailing culture of technological innovation. This has, in turn, led to an evermore technically-literate educational consumer starting to demand new and alternative methods of both accessing and constructing meaning from the educational experience including assessment. As noted by Lonsdale and McCurry (2004), the full range of literacies that

learners have 'need to be valued' (p. 43) and thus there may be a requirement to recognise the differential acquisition of skills across the various literacies. Lonsdale and McCurry (2004) provide an example of this where they assert that, an individual may have limited reading and writing skills, but they may be highly literate when it comes to reading visual images and codes. Conversely, a person may be highly literate when it comes to written or oral communication but they might struggle to use the symbols and metaphors of information and communication technology.

Important in this evolution, particularly within the vocational education field, has been a drive for nationally recognised and accredited competency standards, designed to ensure that the quality and outcome of education and training will be consistent on a nationwide basis. However, in as much as the importance of nationally accredited standards of content delivery are acknowledged, at the same time, it is important to ensure that students who gain accreditation are indeed qualified and competent in the performance of the skills or utilisation of the knowledge for which they have been accredited. If, as Herrington and Herrington (1998) contend, 'assessment design should accommodate learning' (p. 306) then it should be possible to establish a means of measuring the degree of suitability that these hopefully better suited tests should be expected to exhibit. In short, the degree or level of authenticity designed within a given assessment activity should be able to be guided by review against a set of pre-established criteria.

If that is the case, then such guidelines should assist in ensuring that the educational experience, (that is the process of knowledge and skills acquisition) resembles, as closely as possible, the workplace experience that the successful student is likely to face on completion of their education. Whilst this desire to ensure consistency between the

education and workplace experience has traditionally been better acknowledged within the vocational sector, the value of authenticity can have an equally relevant application across the education landscape more broadly, including higher education. Thus it is possible that more authentic assessments could be used to establish more realistic, employment-related applications for many areas of the Higher Education curriculum. This would further assist the learning and assessment experience to become more contextually situated, providing the means to learn not just the required academic component of knowledge, but also the, often underlying, skills that are likely to prove integral to future employment success.

Assessment, authenticity and educational technology

The trend in educational design, to seek out situated or real-world applications for the skills and knowledge taught, aims to enhance the degree of relevance for the learner. By recognising and identifying the crossover between the knowledge and skills components and by encouraging learners to apply new knowledge and skills across a range of areas, it is possible to begin to encourage the further enhancement of broader, more capable thinkers and practitioners. In addition, it will then be possible to extend the workplace relevance of the education system, ensuring the provision of qualified practitioners, who have been formally trained in the means and methods by which they can apply their newly learnt knowledge and skills.

To consider those means and methods of assessing these broader outcomes of learning, a definition of assessment itself must first be obtained, with a particular focus on the assessment of learning outcomes. In this regard, a definition of assessment begins with the premise of Gagné, Briggs and Wager (1992) for whom, monitoring the progress

made by students consists of two related functions: (1) knowing what each student is undertaking to learn and (2) knowing how fast and how well each is progressing. In the first instance, they state that it is important to consider what is to be assessed. In this context assessment is viewed as the means by which the acquisition and demonstration of knowledge and/or skills is determined. However, over recent years, particularly within higher education, as opposed to vocational training, there has been a conscious attempt to move away from the assessment of the reproduction of knowledge, in terms purely of its retention and repetition, to a more rigorous attempt to assess the ways in which such knowledge is applied. Thus, in the context of this study, a discussion as to the ways in which knowledge might be acquired, whether from a theoretical or practical perspective, is key to determining a workable definition of assessment.

This consideration of the role of knowledge in turn leads to a need to determine what it is within any knowledge base held by a student that is actually to be assessed. This is what Gagné, Briggs and Wager (1992) describe as their 'Concept of Mastery' (p. 309). More specifically, they identify the mastery of instructionally-designed learning outcomes as that component of the educational process that should be tested/examined/assessed/reviewed to ensure some form of completion. From a behavioural construct, this process performs two functions. Firstly, it provides the learner with feedback, thereby allowing him or her to acknowledge successful completion of a specific area of training. Secondly, it allows the teacher to audit the success of the training interaction to a given point, and if not yet at the pre-determined level, to intervene in the process, with some form of remedial action. Assessment, in this context, is then a measure of an individual's acquisition of knowledge (declarative, procedural and/or conditional) and a demonstration of mastery of skills (in accordance with a pre-determined learning outcome).

As an activity, however, formal assessment, for the purpose of recording a grade or competence must adhere to certain standards, in order that its outcome can be upheld as a worthwhile part of any learning or training process. Assessment must seek to be valid, reliable, flexible and fair. It must also provide an authentic examination of the learning outcome and be sufficiently rigorous to effectively examine the acquisition of that learning outcome. Assessment should also be current, particularly where it is a test to determine a vocational learning outcome, and ensure the consistent achievement of a specific pre-determined standard.

In this more behavioural educational paradigm pre-defined knowledge constructs are delivered progressively to students, and assessment occurs by means of observation of resultant behavioural change. However, education has seen an evolution from this paradigm to a more constructivist philosophy, which views education as a process in which a student is enabled to construct meaning in their own context from the range of supporting information and content that may be made available. For Herrington and Standen (2000) learning then becomes an active process rather than the result of the transmission of knowledge from program to student. The intention of this theoretical outlook is to recognize that in all environments practitioners actively construct their understanding of a problem and design a sequence of problem-solving steps based on both textbook principles and contextual factors.

The relevance of a constructivist approach to an authentic learning environment is further supported by Wilson (1996) who noted the factors that determine a constructivist learning environment as being ones where ‘learners may work together and support each other as they use a variety of tools and information resources in their guided pursuit of learning goals and problem solving activities’ (p. 5). This

constructivist perspective has impacted upon the perception of the role, value and means of designing and conducting assessment activities. Perkins (1991) asserted that, assessment in education is 'the process of determining whether students have attained curricular goals' (p. 19). This has traditionally meant a focus on retention of knowledge and its application, which in severely limited contexts is measured by what Wiggins (1993, p. 38) described as 'standardised tests'.

The move towards a more constructivist method of teaching and assessing is a reflection of a desire to expand the value of assessment beyond just the measurement of knowledge retention. As well, assessment should seek the means to determine the attainment of higher order educational goals that involve deep understanding and the active use of knowledge in complex, realistic contexts (Herman, Aschbacher, and Winters, 1992). In this context, Reeves and Okey (1996) noted the requirement to design and deliver 'alternative assessment' which for them, 'is absolutely required by constructivist learning environments' (p. 192). In support of this, Reeves and Okey (1996) also noted that 'the very best teachers have used authentic constructivist learning assignments for decades' (p. 192).

This study considered assessment from the constructivist perspective where it was seen as an integral component of the whole learning process and, therefore, a means of establishing more than just the acquisition and retention of knowledge and skills.

Research questions and the study

The problem under examination is that of determining the extent to which authentic assessment, in a constructivist educational context, may provide an effective model for task design and assessment.

The focus of the study is the further refinement of the elements that define authenticity and the construction of a framework to enable designers of educational curricula and content to design assessment activities that will better support the students' application of learning in practice. This could be at the broad stage of curriculum design and development or more specifically in designing lessons and interactions from approved curricula. In order to achieve this, designers may benefit from tools and frameworks to support and guide them in the design of more authentic assessment activity, not the least because assessment design is in itself a complex endeavour.

To address this problem a review and re-design of the final module *Evaluating Educational Multimedia*, of the Australian Army's *Computer Based Learning Practitioners Course* was undertaken. The aim was to bring the module more into line with an authentic approach to learning and assessment. More specifically, the process aimed to ensure that the module's summative assessment activity was actually providing an accurate determination of students' suitability to commence performance of the role in the workplace.

In order to establish whether it is possible to provide such an improvement in student performance, and at the same time inform the broad education field, the question became one of determining whether it is possible to harness the elements of authentic activity to guide the design, development and application of more meaningful, more

authentic assessment activity and thus establish the extent to which authentic assessment could be used to provide an effective model for task design and assessment.

In order to answer this question, it was necessary to address the following questions:

1. What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?

2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?

This study has five main purposes:

1. To establish from the literature the critical characteristics or elements of authentic assessment;

2. To develop those elements into a framework;

3. To utilise expert analysis and feedback to enhance the design of the elements within that framework;

4. To test that framework as a curriculum designer by applying it to the re-design of a module and evaluating the assessment activity from the student perspective; and finally

5. To create learning principles.

The organisation of the thesis

Chapter 2 of the thesis commences with a review of the literature that underpins current thinking on authentic assessment. From the range of definitions of authentic assessment evident in the literature, it was possible to tentatively establish a set of elements that appeared most consistently in the definitions.

Chapter 3 provides a description of the methodology used to determine the elements of this research, in particular, the way in which the draft critical elements were determined. More particularly, the focus of the chapter outlines the way in which these elements were brought together to form an applicable framework. The chapter concludes with some further consideration of the design and development of the learning module upon which the research was based, followed by an outline of the ways in which the module was delivered and its outcomes evaluated.

Chapter 4 sets out the list of the elements determined via the literature review conducted in Chapter 2 and provides the first version of the framework of critical elements describes and discusses the individual elements selected and considers the way in which they were applied in a framework.

Chapter 5 considers how the framework was applied to the design of a discrete module of learning: *Evaluating Educational Multimedia* of the Australian Army's *Computer Based Learning Practitioners Course*. The chapter describes the purpose of the course, and in particular considers the role of the module in the course. The subsequent focus is on the ways in which the elements identified through the literature review were used to design the course and the situated assessment activity. This chapter also seeks to answer the first of the research questions.

The focus of Chapter 6 is upon the implementation of the module within the learning environment and the subsequent analysis of the data obtained in order to answer the second research question. The data collected included that obtained by observation, from written student feedback as well as by interview. This chapter describes the students' experience of the module by means of the collation, analysis and interpretation of students' notes made during its delivery, as well as the observation notes made by the researcher during course delivery, and the video recording that was taken at the same time. Finally in this chapter the analysis of responses to the evaluations sheets completed by the students, as well as the responses given by them during the post course delivery interview is presented.

Chapter 7 provides a summary of the study and outlines its findings. The final chapter, Chapter 8, sets out the elements determined in Phase 4 and the study's conclusions and, considers some of the limitations inherent in this study. It also makes some recommendations for future research.

CHAPTER 2: AUTHENTIC ASSESSMENT:

A GENERAL REVIEW OF THE LITERATURE

Previous findings

This literature review will seek a definition of assessment and consider the role of assessment within the context of higher education. Consideration will be given to the definition of authenticity in assessment and the role and value of the use of information and communication technology within educational assessment, as one of the significant drivers of change in assessment in higher education.

By means of a focus upon specific frameworks, developed and published by other researchers in the field, the literature review will conclude by describing the consistent characteristics or elements considered to be the key determinants of authenticity in assessment.

Assessment

Rowntree (2000) defined assessment in education as occurring:

> Wherever one person, in some kind of interaction, direct or indirect, with another, is conscious of obtaining and interpreting information about the knowledge and understanding, or abilities and attitudes of that other person. To some extent or another it is an attempt to *know* that person. (p. 4)

For Rowntree (2000) assessment was viewed as a 'human encounter' (p. 4) which has six purposes:

- Selection by assessment

- Maintenance of standards
- Motivation of students
- Feedback to students
- Feedback to teachers
- Preparation for life

Ramsden (2004) broadly supported this definition, and further stated that assessment should not be viewed as a linear activity. Rather, it is about a number of simultaneous or linked events, not 'simple dualities such as grading versus diagnosis' (p. 177). For Ramsden (2000) assessment is concerned with a range of factors, including, 'reporting on students' achievements and about teaching them better through expressing to them more clearly the goal for our curricula' (p. 4). It is also about 'measuring student learning and diagnosing misunderstanding in order to help students learn more effectively' (p. 4) This is consistent with Boud (1995) for whom, 'assessment is the most significant prompt for learning' (p. 36).

Crooks (1998, p. vii) provides us with the following eight reasons for assessing:

- Selection and placement
- Motivation
- Focussing learning
- Consolidating and structuring learning
- Guiding and correcting learning
- Determining deadlines to proceed
- Certifying or Grading achievement
- Evaluating teaching

Further to this consideration of the term 'assessment', the review of the literature revealed over the past fifteen to twenty years that there has been a discernable shift towards a constructivist, performance-based, methodology. This, in turn, has impacted upon assessment. As Burke (1997) noted:

> Early theories of learning indicated that educators needed to use a 'building-blocks-of-knowledge' approach whereby students acquired complex higher-order skills by breaking learning down into a series of skills. Every skill had a pre-requisite skill, and is was assumed that after the basic skills were learned, they could be assembled into more complex thinking and insight. (p. xi)

In discussing Bloom's taxonomy, Wiggins (1993) elaborated on the more constructive approach where 'the student must draw upon elements from many sources and put these together into a structure or pattern not clearly there before. His efforts should yield a product.' (p. 215). From the constructivist's view, 'learning is a constructive process in which the learner is building an internal representation of knowledge, a personal interpretation of experience' (Burke, 1997, p. xii). Worthen (1993) developed this notion further. He considered the impact of this change on the role of assessment noting that the majority of the definitions of alternative assessment present two central features: 'First, all are viewed as *alternatives* to traditional multiple-choice, standardised achievement tests; second, all refer to *direct* examination of student *performance* on significant tasks that are relevant to life outside of school' (cited in Burke, 1997, p. xvii). As noted by Wiggins, 'competence is…situational and personal', therefore 'testers should pay most attention to the second of my nine criteria of authenticity that is, replicating or simulating the diverse and rich contexts of performance' (1993, p. 231).

This change in focus for teaching has led to a change in the mindset of many educators, who, when designing assessment activities have had to become what Boud, (1995) describes as 'researchers of students' perceptions, designers of multi-faceted assessment strategies, managers of assessment processes and consultants assisting students in the interpretation of rich information about their learning' (p. 39). This shift in assessment, to an increasing focus upon the demonstration of skills, or the application of knowledge, as components of a more authentic assessment experience, is evidenced by Kerka (1995) who states that, 'assessments are authentic when they have meaning in themselves — when the learning they measure has value beyond the classroom and is meaningful to the learner' (p. 5). Burke (1997) also notes that the emergence of an increasingly authentic emphasis upon assessment design coinciding with 'a push to introduce a variety of assessment methods into our schools which reflect the broader nature of the curriculum' (p. ii).

If, as Jonassen claims (1994) 'constructivism avers that learners construct their own reality or at least interpret it based upon their perceptions of experiences' (p. 34), then, to align with a constructivist perspective, learning must become a, 'constructive process in which the learner is building an internal representation of knowledge, a personal representation of experience' (Burke, 1997, p. xii). This in turn leads to a dilemma for educators, as foreshadowed by Wiggins (1993) who outlined the 'inescapable tension between the challenges presented by contextualized performance and conventional large-scale, generic testing' (p. 206). Wiggins (1993) noted the difficulties inherent in designing tests which can 'better replicate authentic challenges' (p. 210) particularly on a large scale.

Educators also need to consider the placement of assessment as a component of the educational process. Royce (1987) asserted that 'assessment is retrospective, in that it takes account of things already done' (p. 195). Within a more constructivist, situated learning environment, where assessment is embedded, it needs to function not just as a determinant of 'things already done', but also as a means to measure things currently underway.

Educators must also consider the ultimate goals for learning. That is, what is the intended final application of the knowledge or skill for which they are about to facilitate acquisition? Where the intention is to provide learning outcomes that are dependent upon a deeper understanding and application of knowledge, the means and methods of accessing and applying that knowledge should, to a large extent, be informed by the way in which the knowledge will be applied in its final context. As Wiggins (1993) states, 'Understanding is not cued knowledge: performance is never the sum of drills; problems are not exercises; mastery is not achieved by the unthinking use of algorithms' (p. 207). Burke, (1997) further supports this notion stating that: 'New assessments, therefore, should focus not on whether or not students can acquire knowledge, but whether or not they can acquire the disposition to *use* the skills and strategies and apply them appropriately' (p. xii). This is also supported by Puckett and Black (1994) who note that: 'Authentic assessment is not a reflection of inherent capacities, but of individuals' interactions with the environment and their emerging capabilities' (p. 21).

Frohlich (1998) considers the role that emerging electronic educational delivery mediums play in the provision of a more truly flexible, performance-based, approach to educational delivery and assessment. In discussing the potential use of multimodal interfaces, as a means of providing on-going formative assessment of the educational

experience, Frohlich (1998) alluded to a future where 'it should not require the completion and submission of assessment tasks in order to identify that the learner is having problems' (p. 281). Instead, the ongoing monitoring of performance becomes an embedded and integral component of the educational experience. Archbald and Newmann (1988) concur with this understanding of the value of the formative nature of assessment, but also underline the importance of ensuring that the activity itself has a value that is perceived by the student: 'A valid assessment system provides information about the particular tasks on which students succeed or fail, but more important, it also presents task that are worthwhile, significant and meaningful — in short, *authentic.*'

Assessment in higher education

Assessment is, in many respects, the critical component of the educational relationship between student and institution, and within higher education, as with other sectors of the education system, 'motivational benefits are expected to accrue when students can perceive the relevance of learning and assessment activities, thereby enhancing learning outcomes' (Cumming & Maxwell, 1999, p. 177). The challenge for the higher education sector is to be able to demonstrate to students that there is a clear link between the relevance of what they are studying and the ways in which they are assessed. To an extent this is what McLoughlin and Luca (2000) refer to as the 'more pronounced emphasis on the higher education-employment nexus' (p. 2).

'The power to determine the attainment of higher order educational goals that involve deep understanding and active use of knowledge in complex, realistic contexts' (Herman, Aschbacher & Winters, 1992) has become a key driver in alternative or non-traditional means of assessment. One approach to bridging the higher education-

employment nexus has been the application of constructivist theory to both educational delivery and assessment. This approach has enabled the process of assessment itself to become 'a constructivist learning experience, requiring students to applying thinking skills to understand the nature of high quality performances, and to provide feedback to themselves and others' (Rudner & Boston, 1994, p. 3). According to McLoughlin and Luca (2000) 'assessment defines the curriculum and encapsulates the essential learning experience in higher education', and thus the design of educative assessment tasks should perhaps be considered as 'the most important element of tertiary teaching' (p. 1).

In the words of Reeves and Okey (1997), for a constructivist learning environment, 'the focus of the learning activities is on application and active use of knowledge', thus 'assessment in a constructivist learning environments is (and needs to be) as varied and broad as the environments themselves' (p. 195). They identify a number of critical factors which strengthen the authenticity of an assessment. These include: ownership of the task, the fidelity of the assessment and the student's attitude toward the assessment.

In addition, constructivist learning environments, whether open, structured, or virtual, should place learners in positions where they explore, experiment, and actively solve problems (Neimeyer, 1993, p. 93). Thus the design of assessment activities into situated learning environments should, depending upon the domains of learning, seek to provide more authentic assessment platforms. It is to be hoped that it is by the application of such constructivist learning environments into the assessment context that students in higher education will be able to leave their classrooms asking each other, 'what did you learn?' instead of 'what did you get on the test?' (Reeves & Okey, 1997, p. 200).

What is authentic assessment

A more constructivist approach to assessment design has also seen a shift away from what Herrington and Oliver (2000) describe as the 'behavioural science approach to the delivery and assessment of education' (p.23). Within this broadly cognitive perspective, 'learning is concerned not so much with behavioural responses, but rather what learners know and how they acquire it' (Jonassen, 1991, p. 6). Thus constructivism is interested in both 'what learners know and how they come to acquire it'. (Jonassen, 1991, p. 6). The major difference between constructivism and the more objective-influenced behavioural approach is between a view of reality as internally-mediated, as distinct from the objectivist's (behavioural perspective) of reality as being essentially externally-mediated (Jonassen, 1991, p. 8). Thus for the constructivist, Jonassen further asserts 'the emphasis is on how we construct knowledge' (p. 10). The role of assessing the construction of that knowledge becomes an increasingly integral component of the learning activity as opposed to being conducted separately and at the end of it.

The application of constructivism as the prevailing theoretical framework leads into a consideration of the application of assessment within its educational setting. Jonassen (1991) noted the development of 'situated cognition' as the means of ensuring that the appropriate learning environment is constructed to enable the constructive learning experience to occur. But as Jonassen himself warned 'the knowledge that is transmitted may not be the knowledge that is constructed by the learner' (p. 12), and concurrently measured by the assessment.

Thus the move towards this constructivist approach to the design and delivery of education and assessment has led to more detailed consideration of the importance of

the context within which the learning activity is situated. As constructivism places a focus upon the relationship of the learner to the learning context, then the environment, or situation, within which the learning is to occur, becomes an important consideration. As Lave and Wenger (1991) state 'learning is not merely situated in practice — as if it were some independently reifiable process that just happened to be located somewhere; learning is an integral part of generative social practice in the lived-in world' (p. 35).

For Lave and Wenger (1991) this leads to a shift from seeing the 'individual as learner', to that of viewing learning as, 'participation in the social world' (p. 45). This is a shift, 'from the concept of cognitive process to the more encompassing view of social practice' (p. 45). In considering the role of authenticity in assessment as a part of this process, learning seeks to establish 'situated opportunities for improvisational development of new practice' (Lave & Wenger, 1991, p. 97). As Brown, Collins and Duguid (1989) noted 'knowledge, and not just learning is situated' (p. 37). Furthermore, by proposing that 'learning methods that are embedded in authentic situations are not merely useful; they are essential', Brown et al. (1989) affirm the notion of the situatedness of learning. Conversely, 'when authentic activities are transferred to the classroom, their context is inevitably transmuted; they become classroom tasks and part of the school culture', (Brown et al., 1989, p. 34).

In this context, authenticity needs to be considered as a determinant of assessment validity, particularly in terms of the general value of situating an assessment activity within the context of the real world. Despite imperatives to 'embed assessment in real-world contexts' (Custer, 2000, p 3), it is true that assessment often continues to be rooted in outcome more than process. As has been described 'productive, rather than reductive or punitive assessment and accountability systems' have rarely been

developed (Wolf, LeMahieu & Eresh, 1992, p. 9), however, as Wiggins (1993) states: 'We cannot be said to 'understand' something…unless we can employ it wisely, fluently, flexibly and aptly' (p. 207). Thus, 'we should be assessing the student's ability to prepare for and master the various *roles* and situations that competent professionals encounter in their work' (Wiggins, 1993, p. 208), and, at the same time, seek to measure that success in context.

In defining an assessment activity as authentic, an assessment designer is claiming that it will provide a realistic, as well as a valid, determination as to whether or not the learner being assessed has demonstrated the application of the learnt skills and knowledge through the completion of a set activity. Furthermore, if the authenticity of the learning experience, and its assessment, are based upon a situation in a simulation of a real world application, then the degree and context of the reality required from the simulation, (i.e. its fidelity) must also be considered.

From a constructivist perspective, 'knowledge is a function of how the individual creates meaning from his or her own experiences' (Ertmer & Newby, 1993, p. 62), thus, 'learners do not transfer knowledge from the external world into their memories; rather they build personal interpretations of the world based on individual experiences and interactions' (Ertmer & Newby, 1993, p. 63). Authenticity then could be used to provide one reliable indicator of the level of transfer applicable from a summative assessment performance to the workplace. If it is the case that it is the specific interaction between the variables of learner and environment that creates knowledge, then it becomes 'essential that content knowledge is embedded into the situation from which it is used' (Ertmer & Newby, 1993, p. 63),

Thus the constructivist approach assumes that transfer can be facilitated by involvement in authentic tasks anchored in meaningful contexts, and as a result, 'the authenticity of the experience becomes critical to the individual's ability to use the ideas' (Ertmer & Newby, 1993, p. 64). If learning does always takes place in a context and, as Ertmer & Newby (1993) state, 'context forms an inexorable link with the knowledge embedded in it' (p.64), then the authentic situation of the learning and assessment experience becomes a critical element of the learning transaction. From a constructivist perspective: 'If learning is decontextualised, there is little hope for transfer to occur' (Ertmer & Newby, 1993, p. 64). Acknowledging that in a constructivist learning environment, knowledge is inextricably linked to both its context as well as the experience of the learner, the challenge for the designer of the constructivist and authentic learning experience becomes 'to align and design experiences for the learner so that authentic, relevant contexts can be experienced' (Ertmer & Newby, 1993, p. 66).

Whilst authentic assessment has often proved to be more difficult to effect within the more traditional classroom-based educational setting than say the vocational training environment, advances in technology, and its use to produce authentic, if simulated, learning experiences, have gone some way towards closing this gap. Herrington and Oliver (2000) note: 'There is increasing agreement, nonetheless, that computer-based representation and 'microworlds' do provide a powerful and acceptable vehicle for the critical characteristics of the traditional apprenticeship to be located in the classroom environment' (p. 24).

If authenticity is an important aspect of assessment design, then consideration should be given to the role that educational technologies can have in generating assessment

environments or experiences that can be regarded as authentic representations of real world situations.

Assessment and educational technology

Technology as a vehicle from which to access education and training has often followed a well-trodden route to acceptance and general usage. This pathway to wider cultural acceptance masks the real challenge to consider how we may best incorporate future technologies into the design of assessment.

In a brief synopsis of the history of instructional media from 1900 onwards, Reiser (2001) noted the role of instructional media as the primary means of delivering education, as opposed to its often more traditional supplementary role. For much of the past 100 years teachers and textbooks have 'generally been regarded as the primary means of presenting instruction' (p. 54). That is not to say that the application of technology to the field of education and, in addition, its increasing use in assessment has been without its proponents and champions. As far back as 1915, Thomas Edison was informing us that 'books will soon be obsolete in the schools…it is possible to teach every branch of human knowledge with the motion picture' (Reiser, 2001, p. 55).

Of note here is the change in perception of instructional technology, from a supplementary medium for educational delivery to a primary means of delivering education including authentic assessment. With this change in focus has come increased awareness of the diversity of the role of educational technology. McLoughlin and Luca (2000) note 'as institutions move increasingly to online delivery, there is scope for technology to support authentic assessment practices in online delivery' (p. 1). Rather than being seen purely as a means of information provision, it can now be viewed as a

medium that enables and captures learner's ability to apply knowledge and skills in performance. This increases the potential of such technologies as a means of enabling the measurement of performance as a vital component of an authentic assessment.

Reiser (2001) describes the link between instructional design and technology in terms of its positive impact upon the ability to evaluate an improved educational performance. This also provides an additional insight into the desire of seeking to embed an assessment activity into a performance: 'The field of instructional design and technology encompasses the analysis of learning and performance problems, and the design, development, implementation, evaluation and management of instructional and non-instructional processes and resources intended to improve learning and performance in a variety of settings, particularly educational institutions and the workplace' (p.57).

A principal advantage of the use of educational technology is in its ability to provide simulated, interactive and integrated learning environments. It also enables the learning and assessment designer to better integrate assessment activity, both formative and summative, into the learning activity itself. In terms of the authentic application of assessment, this integrated assessment can allow for students to be assessed as they perform a task and create an authentic product rather than having to undergo a separate testing event.

However, assessment appears in the form of low level interactions such as multiple choice quizzes. The increasingly media-rich potential of information and communication technologies will enable the development of ever more sophisticated and more embedded assessment activities.

In the broader discussion as to what technology may or may not be able to assess, consideration should also be given to how successfully multimedia may embed authentic assessment as an integral component of a learning and assessment activity. As Reeves and Okey (1997) note, 'assessments, in constructivist learning environments are as varied and broad as the environments themselves' (p. 195).

In addition to the role that educational technologies can play in the integration of authentic assessments within the learning event, they also have a significant role to play in addressing different learning styles. As Frohlich (1998) notes 'with the emergence of New Media Technologies we may finally be taking our first steps towards allowing learners' preferred learning styles to prevail in the acquisition of knowledge, and its subsequent assessment' (p.281). Of further significance is the development of multimodal interfaces and the potential of such tools to provide for a vast array of new elements in human/computer interaction. These can provide future educational-multimedia software with the ability to recognise specific learners and their individual learning preferences, and also to recognize problems with learning. It should then be possible to design on-going and authentic assessment tasks that can be delivered in real-time, thus reducing the need for formal assessment of the learning at the conclusion of a unit of study. As Dede notes 'like Alice walking through the looking glass, learners can immerse themselves in distributed, synthetic environments, becoming avatars who vicariously collaborate and learn-by-doing, using virtual artefacts to construct knowledge' (1997, p. 165). At the same time, learners can be assessed in a seamless, authentic and integrated manner.

Thus, such technology can offer a variety of ways to support the design and delivery of assessments. At the most simplistic level, it can mimic the models employed within a

more traditional context such as question and answer (with feedback provided), multiple choice questions, or yes/no and true/false, which focus mainly on the assessment of the knowledge retained. More importantly, it can also be used to ascertain the acquisition of specific skills, knowledge or attitudes by means of simulating their practice. Though as Haney and Madaus point out, 'to date, technology has perhaps had a bigger role in improving the power and efficiency of traditional testing than it has in enabling alternative or authentic assessments' (1989, p. 684).

In describing the methods currently in use in the provision of learning and assessment, including via the use of educational multimedia, a brief overview has been provided of the current role and value of the use of multimedia to the consideration of authenticity in assessment. It is also worth noting that as in all forms of assessment design, the underlying assessment principles of validity, reliability and transparency must be upheld, again noting that there will always be a 'tension between reliability and validity in any form of assessment' (Reeves & Okey, 1997, p. 198).

Harper and Hedberg (1997) in describing a number of technology supported learning environments, such as *'Investigating Lake Iluka'* and *'Exploring the Nardoo'*, further elaborate upon the use of educational multimedia courseware that has a design based upon a constructivist approach. A constructivist approach can allow for a focus on the use of cognitive support tools, as a means of helping learners to solve complex or ill-structured problems. Noteworthy, in the consideration of assessment using such tools, is their focus upon the need to be able to demonstrate advanced levels of problem-solving skills, and the importance of being able to analyse and synthesise information. They further describe the importance of recognising higher order thinking skill, as realised through the on-going application of knowledge, as opposed to simply retaining it, in

support of this, Reeves and Okey (1997) describe, 'constructivism demands new approaches to assessment' (p. 192). Duffy and Cunningham further elaborate noting that, 'as knowledge is context dependent, learning [and assessment] should occur in contexts to which it is relevant' (1996, p. 3).

So for future designers of authentic assessment it may no longer be sufficient to rely on the traditional, often external, assessments that concentrate on recall and application of knowledge by means of an external final test. In the development of effective and authentic assessment tasks, whether electronically mediated or not, the designer has to account for not only the impact of task design upon the learner, but also be aware of the broader understandings of what it is that he or she is actually attempting to assess. This is consistent with Reeves and Okey (1997) who note that, 'the ownership of the task is a major factor in strengthening the authenticity of the assessment' (p. 193). Furthermore, they provide a list of the functions that technology should enable in the application of alternative assessment:

> Support for extended authentic activities, the increased portability and accessibility of work, the increased ability to replay performance, the ability to provide libraries or repositories of exemplar performance, the expansion of community participation in assessment and the increased ability to publish and share student work. (Reeves & Okey, 1997, p. 193)

In further support of this is Dede's (1997) notion of 'distributed simulation' which offers students the ability to apply abstract knowledge by ensuring that education is situated in authentic, if virtual, contexts.

Dowsing and Long (1997) further support this notion when they emphasise the importance of being able to 'assess either by means of examining candidate feedback or

by tracking candidate progress through the assessment package'. For them, it is not just the ability to provide correct answers to specific questions that is important. It is also the ability to track and measure the demonstration, by the assessed student, of the appropriate application of the knowledge, skills or attitude and the ability to track their movement through a learning event. Another key advantage of educational technology is that any use of a mouse or other navigation device by a learner can provide a track able map which can be used as a diagnostic tool to analyse trainee performance. This can in turn, be used to enhance the authenticity of appropriately designed multimedia for assessment, by further minimising a requirement for teacher interjection into the learning process.

Information and communication technology mediated education can play an ever important role in the creation of assessment activities that offer increasingly realistic simulations of real world situations. The application of multimedia technology can give the student an opportunity to engage within an authentic representation of the real world. It is via this ability to generate accurate depictions of real world situations that the educational and assessment designer is offered increased opportunities to develop higher levels of fidelity in simulation which in turn can actually enable authentic assessment to become increasingly situated within the performance. As Dede (1997) informs us, 'advances in interface technology have effectively enabled physical immersion in artificial realities that can be designed to enhance learning' (p. 171).

Computer simulations provide an opportunity to bring elements of authentic practice into the classroom, particularly where the assessment activity is delivered via a multimedia tool. However, it does not necessarily follow that all aspects of the generation and collection of the assessment outcomes must necessarily arise from

within the multimedia environment itself. Prestidge and Glaser (2000) in the development of individual assessment opportunities from the group approach to the use of multimedia courseware, offer a range of alternative means of gathering evidence for individual assessment. In particular, they developed rubrics, as a means to represent a standard of performance. By using these rubrics the assessor was able to 'compare a learner's response with clearly articulated criteria for success' (Moon & Callahan, 2001, p. 49). Thus, whilst the learning or knowledge/skill acquisition activity occurred within the electronically mediated environment, its measurement and subsequent analysis was able to occur outside of it.

In support of this is the work of Scott (2000), for whom, 'the key to any effective assessment of performance, live or otherwise, is establishing the criteria and performance indicators in advance' (p. 40). In the context of authentic assessment it is important to know what the determinants of authenticity are, irrespective of delivery mediums. It is not until the determination of the authenticity of performance has been made that consideration can be given to the most appropriate method to assess it. Scott (2000) in reporting the outcome of the *National Center for Research in Vocational Education* study, in particular the use of alternative assessment in vocational education, identified four categories of assessment widely used in alternative vocational assessment: 'the use of written assessments, performance tasks, senior projects and portfolios' (p. 33). However, he also identified the importance of matching the 'appropriate assessment tool to the given learning outcome' (p. 41). Mabry describes this when recognising the importance of, 'matching purpose or outcome expectations with assessment strategies' (1999, p. 125).

Irrespective of the method used to mediate the assessment, it is important to note, as Moon and Callahan state (2001) that the method selected will in itself 'send strong messages about what is important' (p.50). The significant advantages in the use of technology based solutions to address authentic assessment issues are, not only that they can situate the assessment activity in a more authentic context, but that they may also allow for greater flexibility in delivery, to suit the demands of both the student and the educator.

Also relevant is the consideration of the cultural challenge that often occurs with the lack of acceptance of such new technology from the wider educational community. Whilst it is acknowledged that educators have a role to play in what English (2000) describes as 'facilitating the active engagement of learners' (p. 1), it is equally true that the responsibility extends beyond the simple teacher/learner transaction into the learner/assessor relationship. In the same way that it is a part of the role of the teacher to inform and educate the learner into the methodology by which they will learn and be assessed; it is equally, the responsibility of the instructional designer to inform and educate the teacher into the requirements that have been designed into the assessment activity. When this does not occur, anxieties can arise for learners in the application of assessment activities, over and above those that reflect more traditional design considerations. These in turn can impact upon the ability or any assessment tool to deliver on its stated assessment aim.

Finally, the role of educational technology in delivering authentic assessment activities is complicated by the fact that new technologies are often grounded in new models of instructional and assessment design. For example, in describing the use of educational technology to address real world learning and assessment requirements within a number

of United States Police Departments, McCormack (2000) provides some insights into the ways in which an educational activity could be designed to embed realistic, performance-based, assessment from the outset. To train officers in the handling of incidents with potentially armed offenders, the departments use a simulator which can actually shoot the trainee with a small pellet, if they incorrectly analyse the situation and make a consequent incorrect decision as to whether or not to open fire. Such embedded modes of assessment will ultimately seek to not only reinforce the linking of what is taught with the skills required in the workplace, but further enhance the importance of such skills in the mind of the student by increasing the degree of authenticity present in the assessment activity.

Characteristics of authentic assessment

A number of related studies have suggested criteria, many comprised of various critical elements of what may be considered to be determinants of authenticity in assessment. Chief amongst these is the early work in the field of authenticity in assessment by Grant Wiggins (1993) who offered the nine criteria, summarised below, that may be applied in judging the authenticity of an assessment. According to Wiggins an authentic assessment should be:

1. Engaging and worthy problems or questions of importance in which students must use knowledge to fashion performances effectively and creatively.

2. Faithful representation of the contexts facing workers in a field of study or in the real life test of adult life.

3. Non-routine and multistage tasks — in other words, real problems.

4. Tasks that require the student to produce a quality product and/or performance.

5. Transparent and demystified criteria and standards.

6. Interactions between assessor and assessee.

7. Involve response-contingent challenges where the effect of both process and product/performance determines the quality of the result.

8. Trained assessor judgement in relation to clear and appropriate criteria.

9. The search for patterns of response in diverse settings. (Wiggins, 1993, p. 229–230)

In addition to Wiggin's criteria, Herrington and Herrington (1998, p. 308), have proposed 'seven essential elements of authenticity in assessment', within four categories, namely:

Context:

- Requires fidelity of context to reflect the conditions under which the performance will occur (rather than contrived or decontextualised conditions) (Meyer, 1992; Reeves and Okey, 1996; and Wiggins, 1993)

Student's role

- Requires students to be effective performers with acquired knowledge, and to craft polished, performances or products (Wiggins, 1990; Wiggins 1993; Wiggins, 1989)
- Requires significant student time and effort in collaboration with other (Linn, Baker, & Dunbar, 1991; Kroll, Masingila & Mau, 1992)

Authentic activity

- Involves complex, ill structured challenges that require judgement, and a full array of tasks (Wiggins, 1990; 1993; 1989; Linn, et al., 1991; Torrance, 1995)
- Requires the assessment to be seamlessly integrated with the activity (Reeves & Okey, 1996; Young, 1995)

Indicators

- Provides multiple indicators of learning (Lajoie, 1991; Linn et al., 1991)
- Achieves validity and reliability with the appropriate criteria for scoring varied products (Wiggins, 1990; Lajoie, 1991; Resnick & Resnick, 1992)

Using these criteria, Herrington and Herrington (1998) sought to incorporate the elements into a multimedia learning environment in order to investigate the effectiveness of authentic assessment. Whilst they suggested that authentic assessment could be used successfully within an interactive multimedia learning environment, they

did not seek to determine whether the relative degree of authenticity could be considered as a factor in determining outcome.

Further criteria for authentic assessment are provided by Archbald and Newman (1992, p. 72–74), who asserted that, ‘achievement tasks should meet at least three criteria: ‘disciplined inquiry, integration of knowledge, and value beyond evaluation.’ Newman and Archbald (1992, p. 72–74), they also considered that authentic achievement has several characteristics:

- production of knowledge
- disciplined enquiry dependent upon — a prior knowledge
- in-depth understanding
- integration (that is integrating and synthesising knowledge in new ways) and
- value beyond assessment

Cumming and Maxwell (1999, p. 180), outlined four major determinants used in the interpretation of an authentic assessment, namely:

- performance and performance assessment,
- situated learning and situated assessment,
- complexity of expertise and
- problem-based assessment and competence-based assessment

Kendle and Northcote (2001, p. 921), also considered those factors that should be considered as crucial determinants of authenticity in assessment, and developed a series of questions which asked whether the task:

- Necessitated quantitative or qualitative responses?
- Had a clear purpose and outcome?
- Modelled an authentic situation?
- Emphasised process over product?
- Ensured collaborative communications?
- Gave students choices?
- Linked to unit learning outcomes?
- Included feedback mechanisms?
- Encouraged the appropriate discriminatory use of online resources? and,
- Enabled students to examine and present many viewpoints?

Further to this, Wiggins (1993, p. 229-230) added communication as an aspect of authenticity in assessments which 'better replicate authentic challenges and conditions, instead of isolated drill exercises, and draws a distinction as to, 'how a performance or understanding differs from a test of knowledge'.

A final contribution to the discussion on elements of authenticity is provided by Wiggins (1993, p. 234) who in outlining the role of performance in authentic assessment, noted four kinds of constraints facing any performer, namely:

- Demands placed upon us by others
- Limits on the time available to complete the task
- Limits on the human and material resources at our disposal, and
- Limits on our ability to get guidance and feedback as we proceed.

Any meaningfully authentic assessment must ensure that it is cognisant of these constraints, and at the same time take care to ensure that it is not just providing what Cumming and Maxell (1999, p. 188) describe as the camouflage that occurs, when a traditional assessment is 'dressed-up', to appear authentic, often by the introduction of purported real-world elements.

In Chapter 4, I will seek to further analyse and synthesise these elements and refine them into a framework to be used in the design and implementation of a specific module of learning. The trial of this module and the analysis of the outcomes of the trial will be used to determine whether the development of such a framework of critical elements is a viable, measurable and achievable outcome. It is intended that such a framework will ultimately be applied by the designers of both education and assessment.

The next chapter, Chapter 3, will provide a description of the methodology used to determine the outcomes of this research. Particular attention will be given to the ways in which the critical elements are determined and how they will be brought together into a single applicable framework.

CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY

Introduction

This chapter describes the research approach and methodology used to conduct the study. It begins with a description of design-based research that informed the theoretical framework. This is followed by a description of the methods used to collect and analyse the data required to answer the research questions.

Design-based research

Design-based research, sometimes referred to as *development or design research* (van den Akker, 2006), *design experiments* (Brown, 1992; Collins, 1992), or *formative research* (Newman, 1990) has a number of defining characteristics. These include:

- Addressing complex problems in real contexts in collaboration with practitioners;
- Integrating known and hypothetical design principles with technological affordances to render plausible solutions to these complex problems;
- Conducting rigorous and reflective inquiry to test and refine innovative learning environments;
- Defining new design principles.

(Reeves 2006, p. 58)

Van den Akker (1999, p.5) notes that the purpose of a design-based research activity is to reduce 'uncertainty in decision making in designing and developing educational interventions'. This translates to suggesting ways of optimizing the quality of the

educational interventions, thus assisting in the design and development of principles that can be applied, tested and evaluated, and improved. Reeves (2000, p.19) describes 'a growing demand for educational researchers to be more relevant to practitioners such as teachers and corporate trainers'. It is the requirement for relevance and the applicability of the research outcomes to the improvement of practice that determined the application of the design-based research approach to this study.

As Reeves (2000, p. 20) noted on the use of basic research in the field of education 'the value of basic research in such a practical field is limited and that research should have direct and clear implications for practice'. Stokes (1997, p. 99) had also called for the application of an increased 'use-inspired basic research'. In the words of van den Akker (1999, p. 9) 'the major knowledge to be gained from development research is in the form of design principles to support designers in their task'.

Design-based research is distinguished from more traditional design, development and evaluation approaches, as a professional activity. As van den Akker (1999, p. 7) notes design-based research is often initiated 'for complex, innovative tasks'. In the context of this approach and for the purposes of this study, an iterative approach was taken leading to the development of a range of prototypes that 'increasingly meet the innovative aspirations and requirements'. As van den Akker (1999, p.8) further notes 'an iterative process of successive approximations or evolutionary prototyping of the ideal intervention is desirable'.

To be in accordance with this methodology, a comprehensive initial review of the problem under consideration should be made by means of, for example, literature review and expert consultation. Subsequently, efforts should be made to embed theory

into design decisions in order to increase the transparency of the decisions made. However, in utilising such an approach, decisions have to be made concerning the nature of the development-based activities to be researched, and how the research process can be pursued as integral to the process of designing and developing educational content.

Within this approach, Reeves (2000, p. 20), outlined what he considers to be the three major problems with educational technology research. He identifies firstly, a lack of understanding amongst educational technologists about the differences between basic and applied research; secondly, the generally poor quality of published research in the field and; thirdly, the fact that the research often leads to either insufficient or confusing guidelines for practitioners to use to seek to improve practice.

In Figure 3.1, Reeves (2006) conceptualises the design-based research process applied to the problem under consideration.

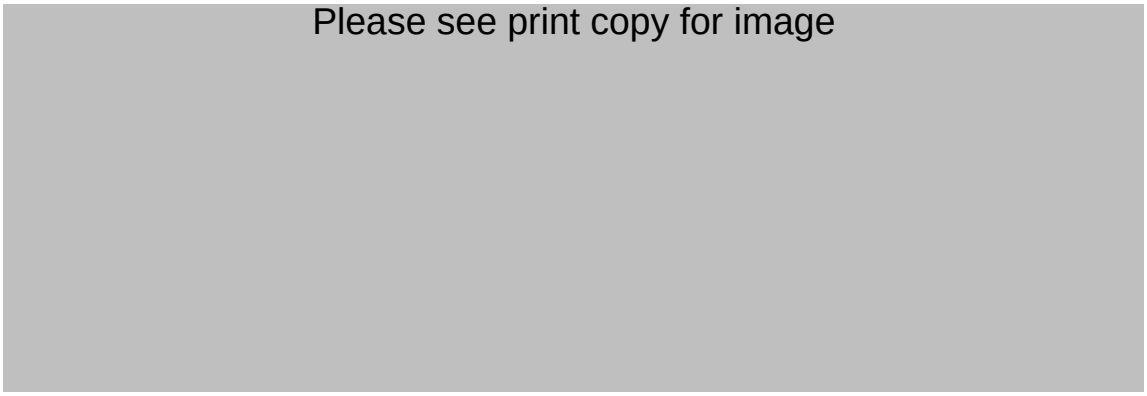


Figure 3.1: Design-based research (2006, p. 59)

He notes that this approach is also supported by van den Akker (2006, p. 58) who states that ‘more than most other research approaches, design-based research aims at making

both practical and scientific contributions'. Estes and Clark (1999, p. 12), describe the application of the process of generic technology development as:

> First identification of the problem, secondly, identification of the relevant research-based knowledge about the problem, and the design of a solution by utilising the research material, thirdly, the need to *package* this solution in a manner that will make it of value to those who will use it and finally, the evaluation of the proposed solution to determine whether it does in fact answer the purpose for which it is intended. (Estes and Clark, 1999)

Further to this, Estes and Clark (1999, p.7) note that 'in applying design-based research to educational technology, the fundamental purpose of science is to generate new knowledge, while the fundamental purpose of technology is to solve practical problems, using whatever knowledge is available and useful'. For the purposes of this study, the appeal of this approach was that it enabled the establishment of design principles applicable to resolving a practical educational problem.

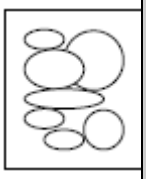
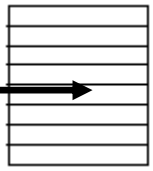
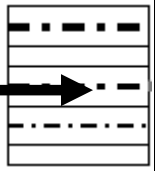
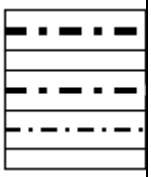
Reeves (2000, p. 24) supports this notion when he states that 'if educational technologists want to be more socially responsible, they should pursue development goals'. Thus, the intention of this study is to determine an outcome with an applicable value, whilst remaining aware of the enduring requirement of the educational technologist engaged in design-based research. Reeves (2000) describe such researchers as 'reflective and humble', acknowledging that the proposed, 'designs and innovations are tentative even in the best of situations' (p.25). Reeves' statement confirms that the overall goal of the design based research approach is to solve what we might consider to be practical problems, whilst simultaneously seeking to develop design principles that can assist with any future decision making within the field.

This is a theme that is consistent with the approach of van den Akker (1999, p. 4) who further describes the increasingly prominent role of design research in the fields of educational technology and media. In particular, he cites Flagg (1990, by van den Akker, 1999, p. 4) who outlines the importance of formative evaluation within media development as a means of seeking to continuously improve a program. This also is consistent with the research model proposed by Reeves (Figure 3.1) and adopted for this study, where the evaluation and testing of a solution in practice, is deemed as fundamental to the design-based research process. Van den Akker (1999, p.4), further highlights the importance of design-based research in the fields of learning and instruction, where, similarly with those of educational technology and media, there is a, 'constant and on-going requirement to seek further refinement of the experience'. In the case of learning and instruction that would include, the design of improved and enhanced learning environments, the iterative formulation of curricula and the design of appropriate assessments to determine cognition and learning.

Finally, Reeves (2000) listed six major types of research goals commonly pursued by educational technology researchers: 'theoretical, empirical, interpretivist, post modern, action and for our purposes, developmental goals'. As Reeves (2000, p. 23) further notes, 'researchers with development goals are focussed on the dual objectives of developing the creative approaches to solving human teaching learning and performance problems while at the same time constructing a body of design principles that can guide future development efforts'. This approach is consistent with the intent of this study, that is, firstly to define and investigate a problem, secondly to design a theoretically-based solution, and then, finally, by means of an iterative design process, implement and evaluate the solution in practice and thus further develop the principles that can be used as a guide to future development.

Table 3.1: The way in which the stages of the design-based research process are applied in this study.

Stages of the design based research process in this study

Chapter	***1**	**2**	***3**	**4**	**4**	**5**	**6**	***7**	**8**
Data		General literature to explore the problem		Discrete body of literature relevant to the defined problem	Interviews	Analysis of initial module assessment and implications for change based on the elements	Student observations Interviews Video material for analysis		
Process		Exploration of the problem Discussion with practitioners		Data mining of specific relevant studies and practitioner feedback	Further practitioner consultation and expert review	Use of expert reviewed elements to re-design module assessment	Implementation and evaluation of the module		Presentation and dissemination of key ideas
Framework of elements (the evolving product)		Key concepts – no elements 		Draft elements 	Elements reviewed by experts 	As before, now applied by researcher 	Modified elements based on evaluation 		Published dimensions 
Framework version				Version 1 (see Table 4.2)	Version 2 (see Table 4.3)		Version 3 (see Table 7.2)		
Phase in Reeves' design based research approach (see Table 3.1)		Phase 1: Analysis of practical problems by researchers and practitioners in collaboration		Phase 2: Development of solutions informed by existing design principles and technological innovations			Phase 3: Iterative cycles of testing and refinement of solutions in practice		Phase 4: Reflections to produce revised design elements and enhance solution implementation
Researcher phases aligned to design based research		Phase 1: Exploration of the problem		Phase 2: Development of a solution			Phase 3: Implementation and evaluation		Phase 4: Presentation of findings
				Phase 2.1: Development of draft elements to guide a solution to the problem	Phase 2.2: Further practitioner consultation and expert review of the draft elements	Phase 2.3: The application of the elements in the re-design of a learning module			

Chapter *1 – Introduces and frames the study, Chapter *3 describes research methodology and Chapter *7 discusses the answers to the research questions

The plan for this research follows the design-based research approach conceptualised by Reeves (2006) using a qualitative methodology, as illustrated above in Figure 3.2. The research process occurs in four inter-related and iterative phases.

PHASE 1: Exploration of the problem

The initial exploration of the problem, which is the focus of this study, required analysis of the literature and consultation with practitioners and a review of current literature was undertaken to determine the ways in which the current designers had been able to gauge the degree of authenticity within an assessment activity and to establish the elements that determine authenticity in assessment. These were then developed into a set of elements that could be applied to the design and implementation of assessment tasks to provide a solution to the problem of the validity of authentic assessment activities. Initial research identified that a number of educators had offered a range of elements to consider depending on the context. It was decided that it was necessary to identify the most consistent of these individual factors and to synthesize them into a framework for the context of this problem. This was undertaken in Phase 2 of the research.

Practitioner feedback

The analysis of the problem that is required by the design-based research process is based not only on an extensive literature review, but also an exploration of the problem from the practitioner's point of view. To this end, a series of conversations and discussions with a numbers of colleagues and fellow educators from both the defence and civilian sectors was held over a period of several months. In all, 13 practioners (three teachers, one course coordinator, four tutors, and five instructional designers) were consulted for their views and insights into the problem. These conversations were exploratory and they were not recorded, but extensive notes were kept by the researcher on the nature and extent of their problems. Each acknowledged the evolving educational importance of finding ways in which to embed the practical application of the subject

matter that they were teaching, coordinating or designing. At the same time, some expressed concern as to the degree of validity in the assessment activities that they were developing to determine often more situated or performance-based, educational outcomes.

PHASE 2: Development of a solution

Phase 2 of the research consisted of three key activities:

1. The development of draft elements to guide a solution to the problem

2. Further practitioner consultation and expert review of the draft elements, and

3. The application of the elements in the re-design of a module.

Each of these activities is described in more detail below.

Phase 2.1 The development of draft elements to guide a solution to the problem

After the initial literature review was conducted to explore the problem in Phase 1, the literature was once again reviewed in more detail. This time the focus was to identify those research studies that had nominated lists of principles and relevant criteria to address the issue of authenticity in assessment.

The approach to the review, a grounded approach, commenced with a broad reading of the literature. Firstly, frameworks currently in existence and then to establish the individual elements that comprised them. Thus, the factors that had been identified to describe whether an assessment may be considered as authentic evolved through a process described by Paton (1980) as:

> I begin by reading through [the available literature]…making comments in the margins or even attaching pieces of paper with staples or paper clips that contain my notions about what I can do with different parts of the data.
>
> (cited in Kelleher, 1993, p.39)

In this way the first list of elements critical to authentic assessment was determined. These elements were then further developed and applied to a selected assessment activity. The students' responses to this activity, based on the identified principles of authentic assessment, form the main source of data to be analysed in the study.

Using this grounded approach a number of critical elements was determined and ultimately used to construct the initial framework. Refereed papers, research studies and reports were identified from the literature that had addressed a similar or parallel issue in assessment. These papers were reviewed in detail and principles and guidelines from each that related to the problem were listed and grouped. From these groups, a list of eight characteristics was extracted which, after review and reflection on the basis of practitioner feedback, were used, as the eight critical elements to form a guiding framework, that was applied for the commencement of the current study. This literature review is described in detail in Chapter 2.

Phase 2.2 Further practitioner consultation and expert review of the draft elements

Expert review of the critical elements was sought in a number of ways. Firstly, during their initial development, each of the emerging elements was discussed, either by face-to-face discussion or by email contact, with a number of current practitioners in the field of education. Through this iterative process, the original list of critical elements evolved. This process itself occurred over a number of months and by means of this content review and refinement the critical elements were evolved, firstly, from a list of

elements drawn from the literature, and, secondly, into a collection of broadly similar elements that represented a number of the elements considered by previous researchers to be indicators of authenticity.

At the end of this review process the list of elements evolved into the eight that were taken as the starting point for more detailed review and consideration within this study. At this point further discussions were sought with three selected experts in the field who had agreed to conduct a more formal expert review. Each expert reviewer was identified on the basis of their extensive experience in the field of authentic assessment as determined from both representations within the literature, and recommendation by practitioners currently working within the field of assessment design, particularly in the area of educational technology.

Each of the experts was forwarded a set of the critical elements for their consideration and feedback. Simultaneously a request was made to conduct a structured interview with them to obtain their views on the proposed list of critical elements, and suggestions for improving the list. (See Appendix 1 for the Expert Reviewer Interview Questionnaire) Two of the interviews were conducted by telephone, the third occurred face-to face.

Initially, the experts were asked to give their opinion on whether the critical elements that they had been presented with made sense to them when considered overall as a framework. Next, the experts were asked to consider and discuss each of the individual elements in turn. This was to determine whether on an individual basis they reflected critical aspects of authenticity within an assessment. In particular, the experts began to deconstruct each of the elements and consider what it was that they were seeking to

establish, and to reflect upon how this might impact upon the determination of authenticity overall. In completing this activity advice was also sought from the expert reviewers as to the terminology used in the framework.

Following this, the experts were asked for feedback or information that might enhance the suitability or applicability of any of the critical elements. In this respect the experts were being asked whether, from their body of experience, there were any aspects of the framework that were incomplete. They were also asked to reflect upon any elements or areas, currently not included, that might further enhance these elements. Finally, the experts were invited to provide any further comment that they believed could be used to further enhance the quality of either the framework or the selected individual elements. The three experts provided extensive feedback on each of the elements selected from the perspective of both their individual value as well as the value that they might hold as a part of a framework. On completion of this collaborative process, the feedback was collated, analysed and the framework of draft elements was revised in the light of that feedback. The expert review is detailed in Chapter 4.

Phase 2.3 The application of the elements in the re-design of a module

The framework of critical elements was then used to guide the design and development of a solution. A suitable course was required to instantiate the guidelines derived from the literature review, practitioner consultation and expert review. A module of the Australian Army's *Computer Based Learning Practitioner Course*, namely, *Evaluating Educational Multimedia* was selected for this purpose. This module had been delivered to a previous group of students the year before as the final module of the course. On that occasion, although it had been intended as a means of providing students with a series of workplace applicable skills, post course evaluations both at its completion and,

subsequently in the workplace, had identified several problems in its implementation. It had been observed that the theoretical focus of the module, and its reliance on PowerPoint and presentation by the teacher as the main mode of educational delivery, had impacted upon the students' perceived levels of confidence in actually undertaking this task within the workplace. The course had also been scheduled for re-design in the intervening twelve months, with an increased emphasis being placed upon the use of a self-paced, distance delivery approach being introduced, where appropriate, into the course design.

The re-design was based initially on the feedback received from students and this is described in more detail in Chapter 5. In particular, it emphasised not only embedding increased levels of authenticity into the learning experience, but more particularly, incorporating more authentic assessment outcomes as one means that might provide the students with improved confidence in applying their workplace skills.

The design process involved iterative cycles of testing and refinement to the solution in practice supported by information gained by means of informal evaluation and on-going testing of the prototype, during its design and development phase. This was a fundamental component of the study's research design and represented another of the iterative phases of design, reflection and re-design necessary to the application of the selected design-based research approach to this study, which consistent with the acknowledged aims of design-based research sought to address a complex problem in a real context in collaboration with practitioners. It also involved on-going informal review and evaluation of the initial written course materials, including curriculum and design documentation such as course design notes, curriculum documents and lesson plans, followed by the on-going review and evaluation of the learning module as it was

being developed. This process was dependent upon an ability to gain access to fellow educational designers with sufficient knowledge and experience to enable them to contribute to the refinement of the designs and prototypes. The nature of the environment within which the module was re-developed provided good access to such educational design staff and practising teachers. Moreover, the design methodology of the continuous improvement, or iterative re-design, of the course's design in practice was consistent with the learning design and development methodology employed within this environment.

To that end the Army's Training Technology Centre utilised an iterative four stage learning design process comprised of 'analysis', 'design', 'development' and 'test', prior to the implementation in practice of that learning content and the use of student and course evaluation data to review, and where necessary, revise the learning material in practice.

These distinct stages each encourage the opportunity for review at any point in the design or development process, of either the content or structure of that module. In this respect the design and development process applied was itself an iterative one that both enabled and encouraged review and reflection of the proposed design. At the same time, the Training Technology Centre often applied a process of rapid prototyping that further encouraged the on-going review of learning content through the application of the design and development stages. Thus, as described, this process was consistent with the third stage of the Reeves' design based research model (Reeves, 2006, p. 59) where there is a requirement to ensure that, 'iterative cycle of refinement of a solution occurs in practice'.

This final module and its associated assessment activities were specifically selected as being suited for re-development for delivery as a multimedia learning experience to be capable in the longer term of delivery by distance. For the purpose of this trial, the self-paced nature of the distance experience was simulated within a supervised classroom environment.

The revised framework of critical elements was used as the basis for the design and development of the module and its summative assessment, a process that was undertaken in two distinct stages. In the first design stage, the learning module was re-designed using the critical elements framework. During this time the researcher as previously described took design notes and reflected upon the overall design process applied to the iterative design and re-design of the module in consultation with professional colleagues. This included continuous informal feedback from fellow educational designers. Next, during the second development stage, the outcomes of the design stage, (e.g. course guidelines, curriculum documents, planning notes, lesson plans) were used in the development of the final module and, subsequently, gathered for future analysis.

A full description of the design of the learning module and how each of the critical elements from the framework was instantiated in the module is described in more detail in Chapter 5.

PHASE 3: Implementation and evaluation

Once the module had been redesigned and developed according to the draft guidelines, it was implemented and evaluated in practice. The module would be the final module

delivered to the students over a period of two days at the end of an intensive two week residential course.

It was on the basis of the data obtained from this field test that the framework was further refined.

Participants

The intent of the Computer Based Learning Practitioner's course was to train a group of six male (n=4) and two female (n=2) Army Educational Corps Lieutenants and Captains (or Royal Australian Navy equivalents) in the basic principles of designing and developing training courseware for delivery by means of educational multimedia. Each of those panelled and selected for this course had been posted to positions within Army and Navy Training Command as Computer Based Learning Practitioners, and would take up these postings in the year following completion of the course. The purpose of this role within a Defence Force Training Establishment is to provide advice to curriculum developers as well as training providers as to how educational multimedia might be best employed to enhance training. An additional purpose of the role is to manage the design and development of the courses, or modules of courses, selected for delivery by such means. This module had been initially designed with the intent of providing students with a series of workplace applicable skills that they would be able to apply upon return to their military workplaces. However, based on the data from the subsequent evaluations of the students' performance on completion of the previous implementation of the course and in the workplace, it appeared that the module had not enabled students to perform sufficiently. The feedback received from the evaluation of both students and supervisors within the workplace indicated a general feeling that the previous course had been overly reliant on the teaching of theory, within a face to face,

classroom environment, with too little student interaction and insufficient opportunity to apply the theory taught in practice. This had impacted negatively upon students' confidence levels as they had acknowledged at evaluation on course completion.

Each of the participating students who undertook this revised course of study arrived on the course with a variety of knowledge and understanding of the application of educational multimedia in the face to face or distance mediated educational context, though all were qualified and had experience as high school teachers (a pre-requisite for entry into their chosen military career), and each had at least one prior two-year posting within a defence military training environment as a Training Developer. This posting, another pre-requisite for attendance in the Computer Based Learning Developers Course, would have enabled them to build upon their broader understanding of the principles of good educational practice, with the requirement to design and develop the curriculum and curriculum documentation upon which all military training is based. A pre-requisite for a Training Development posting is the successful completion of the Australian Defence Force's *Training Developer Course*. The *Computer Based Learning Practitioners Course* was to be delivered near the end of the calendar year to the six Defence Force students panelled to undertake this training.

The module was delivered as the final component of the course overall. The course itself was a two week residential course, and this module was conducted over the final two days of that course. Whilst other modules of the course had undergone limited re-design, in line with the principle of increasing the amount of the course that could be delivered by distance, it was this final module that had undergone the most significant amount of redesign, and although still conducted within a classroom environment on

this occasion, it was designed and intended to be used subsequently as a self-paced module delivered at distance.

Course delivery context

To undertake this module the students were located in a classroom environment, with individual personal computers arranged around three of the four walls of the room. In addition, each of the students was provided with an individual desk set in the middle of the room facing towards the front of the classroom, in a more traditional classroom design. Each student also brought with them a laptop computer issued at the commencement of the overall two week course. The classroom was situated within a Defence education facility that was mainly comprised of classrooms with an additional limited office space for administration. Co-located with the Education Centre was the local Defence Library for the region. The building and nearby facilities also included a kitchen area that was stocked with tea and coffee making material as well as a fresh water fountain. A covered pergola area was made available at the back of the building, next to the kitchen area for the students' use. The room was set up initially with the desks presented individually in three rows of two desks per row, and each student selected where they wanted to sit. Students were able to move furniture if they wished to better accommodate their work practices.

Students were informed from the commencement of the module that they were able to work their way through the content of this course at a pace that suited them. If they were not able to complete the assessment activity within the two days allocated for the module, they could negotiate with the teacher/researcher to hand in the assessment at a later date. Students were also informed that they would be entitled to use both the

classroom and the area immediately surrounding the classroom during the conduct of the training.

It was intended from the outset that the implementation of the training module would, as far as it was possible, seek to reflect the likely work environment within which the students would subsequently find themselves employed. To this end the student body itself represented a mixture of commissioned officer ranks from Army Lieutenants to Captains and a Royal Australian Navy Lieutenant (Army Captain equivalent). Whilst this did accurately represent the likely range of peers with whom they would need to collaborate within the workplace, it was noted that they were commissioned officers only, and in the true military workplace they would often be expected to supervise and work alongside non-commissioned officers and soldiers, as well as civilian staff. However, the work environment that was established for them could, to all intents and purposes, be considered as being authentic.

As well as being provided with laptop computers for their personal use, students were given access to the classroom computers so that they had broadband access to the internet that they would be likely to need to complete the module. They were also given a large degree of personal freedom when undertaking the module, again to be consistent with the defence office work environment. Thus, students were able to come and go at various times throughout the day, acknowledging the expectation to have completed the training, and to be sure that they were able to either hand in the completed expected outcome to the pre-determined deadline or at a subsequent negotiated time.

However, what did represent a unique experience for the students was the fact that, unlike their normal military work environment, they were not seeking to juggle multiple

tasks and deadlines. Instead, in the training environment, they had the comparatively unique experience, from a work perspective, of only having to fulfil one expectation at a time. This meant that although they were expected to discuss and collaborate with one another in the completion of this activity, they were, by and large, free from the distractions of competing priorities, and the factors that go along with this such as telephone calls and personal interruptions from colleagues often working on other priorities. So, as can be seen, an attempt was made to ensure that a high degree of fidelity was provided within the students' work environment. In addition, the nature of the challenge that confronted the students, that is the requirement to develop a product as the final assessment outcome was also consistent with both the critical elements as well as the workplace. In particular, the design of the module was such that it also required that they be able to demonstrate the ability to successfully transfer knowledge obtained during previous modules of the course, as well as this module and further to this, that they demonstrate an ability to both critically reflect upon the outcomes produced and, in addition, be able to discuss and collaborate with one another as required.

These data collection methods were chosen for a number of reasons, particularly for the opportunity that they presented to gather the data in the learning setting, and the ways they enabled corroboration of data during the subsequent analysis.

Data collection from students

A range of data formats were collected for subsequent analysis for the purpose of this study.

1. Noted observations

Whilst the training was being conducted the students were observed by the researcher and extensive notes were made as to the ways in which they interacted with both the material as well as each other. Subsequently, these notes were collected, collated and analysed.

2. Classroom video

During the training, various aspects of the activity were video-recorded for the purposes of subsequent analysis in order to ensure that noted observations made during the conduct of the activity were consistent with the notes made at the time. In both instances information was obtained as to the ways in which the students interacted with both the learning content well as each other.

The particular aspects of the training video-recorded were those that occurred at times of particular debate or discussion, and also, at times when the students were mobile within the classroom, in order to establish whether any particular patterns could be determined from these interactions

3. Written evaluation questionnaire

At the completion of the training each student was asked to complete a written questionnaire to provide their opinions on a number of issues relating to the impact of authentic design upon the development and delivery of the module. These written responses were collected by the researcher and reviewed prior to a more detailed examination of the information that they contained by means of a follow-up, face to face, one on one, interview conducted with each student.

The questionnaire (see Appendix 2) was intended as a means of establishing the level of understanding that the students themselves had with regard to the implementation of authenticity in education. Further to this they were asked to express their own views as to how well, or otherwise, they considered that the critical elements used in the design of this module had performed in practice. This they did by means of reflecting individually upon each of the elements used within the design process, and commenting how effectively they believed that each had been implemented within the module.

The students next were asked to consider the ways in which the critical elements themselves had been expressed, and to comment on whether they thought that the questions could have been better, more clearly or more concisely phrased in order to clarify their meaning to potential future educational designers who might use them. The students were also requested to comment upon whether they believed that these elements could have been more fully applied to further enhance the authenticity of this assessment activity.

In conclusion, the students were invited to present their views as to whether they thought that there were any issues or flaws in the elements that they had considered or whether any other factors that should have been taken into account were missed.

4. Interviews

The final stage of the data gathering process was that of the interview to confirm, clarify and further explore the written responses received. The questionnaire (see Appendix 3) was used to elicit further clarification on any issues that were not clear from the written feedback received. Greater detail was sought on a number of the issues or concerns that the students raised from their experiences. The students had previously been encouraged

to provide as much detail as they felt able, and were assured that all responses would be dealt with in-confidence. Completion of the written evaluation questionnaire and the subsequent interview, were undertaken entirely on a voluntary basis, but for the purposes of this study each of those involved indicated a willingness to fully participate in all aspects. At the time of the interview, the researcher recorded by hand, in writing, the responses that the students were providing on a blank questionnaire. It should be acknowledged here that the researcher had had significant note taking experience in a range of roles and positions. It was intended that these interviews would represent a mechanism to both confirm the conclusions drawn from the collation of the notes made during observation and also to ensure that these conclusions were consistent with the views that the students themselves held. The collated responses of each of the six participant students are described in more detail on an element by element basis in Chapter 6.

PHASE 4: Presentation of findings

The fourth phase of the research is to consider the extent to which authentic assessment provides an effective model for task design and assessment. This will be achieved by means of the presentation of a final set of critical elements that could be placed into a revised framework based upon the data received at the conclusion of Phase 3. This will be described in Chapter 7. In reflecting upon the relative value of the critical elements used in the design of this module, as determinants of authentic assessment activity, it was important to establish the value of each of those elements in determining the extent to which authentic assessment provided an effective model for task design and assessment.

At the completion of this process the researcher was able to consider the ways in which these specific characteristics of authentic assessment could facilitate the design and assessment of complex and authentic tasks.

Summary of the research plan

This study followed a design-based research approach, as illustrated in Figure 3.1, over four phases, each of which had a separate intent. The rationale of the first phase of the research was to explore the problem by means of the analysis of the literature and consultation with practitioners. In the second phase, the rationale was that of developing a solution by undertaking three key activities. These activities were the development of draft elements to guide a solution to the problem, the obtaining of further practitioner consultation and expert review of these draft elements to further refine them and, finally, the application of the elements in the re-design of a learning module.

The third phase was the implementation of the learning module its evaluation and the collection and analysis of the data that arose from it. This was undertaken in order to investigate the effectiveness of the framework itself, as defined in the second phase, in the provision of an alternative model for the development of tasks in a flexible learning environment. This phase also sought to both isolate the specific design characteristics of the assessment activity, at least in so far as they reflected authentic assessment practice, and to assess both the importance of, and relationship between the defined elements. The fourth and final phase was that of considering the extent to which authenticity provided an effective model for task and assessment design and the development of a final set of critical elements into a revised framework based upon the data received at the conclusion of Phase three.

Table 3.2 provides an explanation of each of these four distinct phases, with an outline of the intent of each phase along with a description of the research question that it supports and a further description of the intent or aim of that question. Finally, this table also sets out the method to be used to establish the necessary information against each of these four phases and the data that would be required for them.

Table 3.2: Summary of the research plan

Phase	*Intent*	*Research Question*	*Aim*	*Method*	*Data Required*
1	Analysis of the problem, Review of the Literature Consultation with practitioners		To determine and define the nature of the problem	Literature review and consultation with practitioners	**Research and theoretical papers on authentic assessment** **Practitioner commentary**
2	Intensive review of literature for principles and guidelines Creation of draft framework Design and development of learning environment		To define the critical elements of an authentic assessment framework and to collect data from fellow practitioners	Literature review and analysis of selected studies Case study based on design and implementation of a module designed to incorporate elements of authentic assessment In-depth analysis of course design documents and artefacts	Research, evaluation and theoretical papers on authentic assessment from the mid 1980s to the present Course design notes, curriculum documents, lesson plans Practitioner feedback and expert review
3	Implementation of the learning environment Data collection and analysis from module participants	To what extent does authentic assessment provide an effective model for task design and assessment?	To investigate and evaluate the effectiveness of the framework of authentic assessment, as defined in Phase 2 of the research, and to collect data from the field to test and refine the framework	Observation of implementation of course Non-scheduled standardised interviews with: • Students Experts/elite practitioners	Observation notes and records, video Interview data

		What are the specific characteristics of authentic assessment that facilitate the design and assessment of complex and authentic tasks?	To isolate the specific design characteristics of an assessment task which enable it to effectively reflect authentic elements of assessment	Case study On-going literature review In-depth analysis of course design documents and artefacts Observation of implementation of course Non-scheduled standardised interviews with: • Students Experts/elite practitioners	Research papers, reports and conference proceedings Course design notes, curriculum documents, lesson plans Observation notes and records, video Interview data
		How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	To assess the importance of, and the relationship between, the elements defined as critical characteristics of authentic assessment	Case study Observation of implementation of course Non-scheduled standardised interviews with: Students	Observation notes and records, video Interview data
4	Development of final guidelines			Use of data obtained at the completion of phase 3 for the revision of the framework	

Conclusion

This chapter has provided a description of the research approach and methodology used to conduct the research. The next chapter, Chapter 4, will describe and discuss the draft individual elements in more detail and set out the way in which they have been applied to the formulation of a framework of critical elements.

CHAPTER 4: AN EFFECTIVE MODEL FOR TASK DESIGN IN FLEXIBLE LEARNING ENVIRONMENTS

This chapter describes the elements considered critical in the determination of authenticity in assessment, and also details the process by which they were evolved. This is a key design process in Phase 2 of the design-based research approach and Phases 2.1 and 2.2 of this research. The description of the process is enhanced with a detailed description of each of the elements selected and a consideration as to how they relate to one another within the framework.

Analysis of the elements of authentic assessment

The review of the literature in Chapter 2 identified a number of researcher perspectives on the scope and interpretation of various elements that might be used to determine authenticity of assessment activity. These elements are analysed to provide a single cogent and applicable framework that will enable their direct application to the design and development of an assessment activity. The list of the elements of authentic assessment established is described below in Table 4.1.

Table 4.1: Researcher's synthesis of the elements of authentic assessment from the literature

Characteristic of authentic assessment	*Explanatory question*	*Supporting literature*
1. Challenge presented	Are the students required to demonstrate the application of skills and knowledge in a selective manner?	Lund (1997), Herrington and Oliver (2000), Moorcroft, Desmarais, Hogan and Berkowitz (2000), Hart (1994, in Moorcroft, Desmarais, Hogan and Berkowitz, 2000), Newmann, Marks and Gamora (1995), Newmann and Wehlage (1993), Scott (2000), Alsami (2001)
2. Product as an assessment outcome	Is successful outcome of the assessment activity determined by means of a measured outcome (product or performance)?	Archibald and Newman (1988, in Peterson, 2002), Lund (1997), Herrington and Oliver (2000), Brown and Craig (2000), Moorcroft, Desmarais, Hogan and Berkowitz (2000), Newmann, Marks and Gamoran (1995), Newmann and Wehlage (1993)
3. Transfer of learning	How closely will the assessment experience replicate the workplace?	Tanner (1997, in McAlister, 1994), Berlak (1992, in Tanner, 2001) Hattie, Biggs and Purdie (1996, in McAlister, 1994x), Tanner (2001), Wolf (1993, in Supovitz and Brennan, 1997), Newmann, Lopez and Bryk (1998, in Avery, 2000)

Characteristic of authentic assessment	*Explanatory question*	*Supporting literature*
4. Critical reflection & self-assessment	Does the activity encourage critical reflection on the meaning beyond the assessment experiences?	McAlister (1994), Newmann and Wehlage (1993), Scott (1994), Khattri, Reeve and Kane (1998, in Hoepfl 2000), Cizek (2000), Herrington and Oliver (2000), Aschbacher (1995, in Moorcroft, Desmarais, Hogan and Berkowitz, 2000), Moorcroft, Desmarais, Hogan and Berkowitz (2000),
5. Accuracy in performance outcome, and fidelity of assessment environment	Are the knowledge/skills being assessed critical to the workplace performance and is it important that the assessment environment is consistent with the workplace?	Newmann and Wehlage (1993), Moorcroft, Desmarais, Hogan and Berkowitz (2000), Tanner (1997, in McAlister, 1994), Herrington and Oliver (2000), Avery (2000), Lund (1997), Newmann and Wehlage (1993 Wiggins (1990)
6. Fidelity of assessment tools	Are the tools used for the assessment activity consistent with those used in the workplace?	Berlak (1992, in Tanner, 2001), Newmann, Lopez and Bryk (1998, in Avery, 2000), Northcote and Kendle (2000), McLellan (1994, in Northcote and Kendle, 2001)
7. Discussion and feedback	Does the assessment activity encourage discussion and feedback?	Newmann and Wehlage (1993, McAlister (1994), Northcote and Kendle (2000), Herrington and Oliver (2000)

Characteristic of authentic assessment	*Explanatory question*	*Supporting literature*
8. Collaboration	Does the assessment activity encourage collaboration?	Northcote and Kendle (2000), Herrington and Oliver (2000)

Practitioner feedback

The elements described in Table 4.1 were reviewed and discussed by thirteen education practitioners as described in Phase 2.2. This process was largely an informal activity and relied mainly on making use of available opportunities to discuss both the concept and value of authenticity as a determinant in valid assessment design, as well as giving consideration to the range of the elements provided. Although largely informal in conduct, the researcher did keep noted records of these conversations, often however, having to write-up these notes in summary as soon as possible on completion of the conversation. The intent of this activity that occurred over a period of approximately one month was, firstly, to test the perceived relevance of the characteristics provided and, secondly, to seek advice as to how best these broader characteristics might be represented in an applicable framework.

Based upon the outcomes of these conversations and feedback, Table 4.2 (below) presents the characteristics or elements of authentic assessment as determined from the literature and discussed with practitioners. It then provides a summary of the feedback received from practitioners with reference to each of these characteristics, and based upon this feedback, describes the way in which the characteristics of authentic

assessment will be represented as a critical element of authentic assessment for further consideration and discussion in the next part of this process, namely Expert Review.

Table 4.2 Researcher's translation of characteristics to critical elements of authentic assessment with practitioner feedback

	Characteristic of authentic assessment from the literature	*Summary of practitioner feedback*	*Critical element of authentic assessment*
1.	Challenge Presented	• This characteristic is appropriate but needs to be more specifically focussed on stating that the challenge is upon that which is provided to the student in completing the activity. • Challenge is a relative concept and it needs to more clearly articulate this by means of identifying the degree or level of the challenge implicit in the assessment activity.	Degree of challenge(s) presented to the assessed student
2.	Product as an assessment outcome	• It is not always the case that an assessment outcome is represented as a product and in many instances it might be that a performance is an equally valid indicator of successful assessment outcome. • It should be acknowledged that in considering the relevance and value of product (or performance) that it is often not until completion of the activity that it is possible to appropriately make judgements about them as indicators of successful assessment outcome.	Performance, or product, as final assessment outcome

	Characteristic of authentic assessment from the literature	*Summary of practitioner feedback*	*Critical element of authentic assessment*
3.	Transfer of learning	• When consideration is given to the role of transfer of learning in determining performance at assessment it should stated that this can apply equally to skills, knowledge and attitude. • It should also be stated more clearly that successful assessment performance will require that some degree of transfer is required.	Transfer of learning (skills, knowledge, attitude) required
4.	Critical reflection and self-assessment	• In terms of addressing the role of critical reflection and self-assessment, it should be clearly stated that a successful assessment outcome will require the student to demonstrate that they have been able to critically reflect upon their performance or product and accurately self-assess them in terms of improving overall assessment outcome	Critical reflection and self-assessment or evaluation required
5.	Accuracy in performance outcome, and fidelity of assessment environment	• The accuracy of the performance (or product) provided by the assessed student is a necessary determinant of successful assessment performance; therefore, it is of value to state them as a requirement. • If accurate performance or products are considered as pre-requisites to successful assessment outcome, then consideration must also be given to the fidelity of the environment in which the assessment activity has occurred. In this respect, it is more difficult to make judgements on either product or performance if there are inconsistent compromises made in the establishment of the environment within which it occurs.	Accuracy in product or performance, and fidelity of assessment environment, is displayed

	Characteristic of authentic assessment from the literature	*Summary of practitioner feedback*	*Critical element of authentic assessment*
6.	Fidelity of assessment tools	• Similarly to the accuracy of the environment within which the assessment activity occurs, so too must the tools used by the student in undertaking the assessment accurately model those that will be applied in the workplace.	Fidelity of assessment tools used
7.	Discussion and feedback	• The ability to be able to discuss a product or performance and be able to receive and distil that feedback and use it to improve both final product or performance are vital to many aspects of successful workplace performance and therefore should be required if the assessment outcome is to be considered valid and reliable.	Discussion and feedback required
8.	Collaboration	• Increasingly successful workplace performance is based upon the ability to work in teams or to be able to lead teams. In this respect, the ability to be able to collaborate with peers in the execution of a successful performance, or the completion of a successful product, should be regarded as an essential requirement for authentic assessment design.	Collaboration required

Evolving and further developing the critical elements

The outcome of this Phase 2.1 was the list of the critical elements shown in Table 4.2 and detailed below:

1. Degree of challenge(s) presented to the assessed student
2. Performance, or product, as final assessment outcome
3. Transfer of learning (skills/knowledge/attitude) required
4. Critical reflection and self-assessment or evaluation required
5. Accuracy in product or performance, and fidelity of assessment environment, is displayed
6. Fidelity of assessment tools used
7. Discussion and feedback required
8. Collaboration required

A detailed description of the individual elements within Table 4.2 is provided below, on an element by element basis:

1. Degree of challenge(s) presented to the assessed student

As Lund notes (1997), authentic assessment tasks establish connections between real world experiences and school-based ideas. They also present students with the full array of tasks that mirror the priorities and challenges found in the best instructional setting. This idea is expanded upon by Herrington and Oliver (2000, p. 23), who list the requirement to provide authentic activities as the second of their nine situated learning design elements.

Thus within an authentic assessment activity, students are required to demonstrate their ability to synthesise, from the range of skills and knowledge that they have acquired, those which will be necessary for the completion of a specific outcome or outcomes. This process requires both analysis of the task, and the subsequent selection of the appropriate response. Degree of challenge is a reflection of the authenticity of real world situations and tasks where the necessary response will often require the synthesis of a range of skills and information into the formulation of a potentially correct response, but the approach to that response may not always be clear cut or obvious.

This notion of the synthesis of skills and knowledge is further borne out by Moorcroft, Desmarais, Hogan and Berkowitz (2000, p. 20), who assert that, 'authentic assessments are designed not only to be assessment tools but also to be exercises through which students explore their understanding of a topic and apply that knowledge'. Students are challenged to utilise, from the array of knowledge and/or skills, that they have acquired those elements that will enable them to provide such an appropriate response. Moorcroft et al. (2000, p. 20), further expand on this notion of skill/knowledge selection within authentic assessments noting that 'the focus is on developing understanding and applying knowledge, rather than assessing achievement alone'.

However, implicit in this notion of developing understanding and applying knowledge is the need to select, link and utilise the appropriate skill and knowledge elements. In his consideration of performance assessments, Hart (1994, p. 4), asserts that 'a good performance assessment…begins to engage students through tasks like…problem solving'. It is this process of engaging through the use of problem solving skills that help to generate the notion of greater authenticity in the assessment process. In a similar vein, Hart describes a process of 'obtaining information through hands-on

experimentation and then synthesising and applying this new information to form a hypothesis or interpretation [or product or outcome]'.

Degree of Challenge is then a critical determinant of authenticity within a given assessment activity, what Newmann, Marks and Gamora (1995, p. 1) refer to as the 'challenge of constructing or producing meaning or knowledge, instead of simply re-producing meaning and knowledge as created by others'. Though, as Newmann and Wehlage (1993, p. 8) note 'students construct meaning and produce knowledge as opposed to just repeating declarative knowledge'. In support of this, is Scott's assertion (2000, p. 31), that, 'skilful and effective teachers require students to analyse and synthesise information'. He further notes (2000, p.31) that assessments should, 'tap the connectedness of concepts and the student's ability to access interrelated chunks'.

Finally Alsami (2001, p. 27) notes that teachers are increasingly expected, 'to be able to teach by applying knowledge and material to the lives of the students'. Thus, students must be expected to demonstrate competency or capability in similar ways by means of reviewing, considering and applying that which is relevant to the completion of a satisfactory assessment outcome.

2. Performance, or product, as final assessment outcome

This element considers the issue of assessment outcome in terms of either a performance or product. The requirement is to create or develop a performance or a product as a part of the final assessment, and, where appropriate, one that is consistent with workplace expectation. Thus, a designer needs to determine the extent to which the assessment activity requires the production of a completed outcome or product. As Archbald and Newman (1988, p. 33) assert 'students demonstrate skills and knowledge

by engaging in complex performance, creating a significant product or accomplishing a complex task using higher order thinking, problem-solving and often creativity'. This notion is further supported by Lund's (1997, p. 25) view that 'authentic assessment requires presentation of worthwhile and or meaningful tasks that are designed to be representative of performance in the field'. The idea of performance or product finds further resonance with Herrington and Oliver (2000, p. 23) who lists the requirement to 'articulate', as a means of ensuring that 'tacit knowledge… [may]…be made explicit'.

It is then by means of the application of such skills and knowledge in the workplace that a crafted outcome is produced whether it is a performance or a product. As Brown and Craig (2004, p. 2) note, 'authentic assessments focus on determining the skills and knowledge that the students are able to demonstrate while completing specific tasks'.

It may be that the actual application of a specific set of skills and knowledge in a particular order may be subservient to the requirement to produce a functional final performance outcome or product. This is further supported by Moorcroft et al,'s (2000, p. 20) view that, 'authentic assessments are designed not only to be assessment tools but also to be exercises through which students explore their understanding of a topic and apply that knowledge'.

It is in this application of knowledge and skills to a topic to produce an outcome that success is often measured. That is, the end may very well justify the means, and provided that an acceptable performance or product is elicited, then the workplace outcome may be considered to have been successful.

Employers can often be reluctant to review the means by which a successful outcome has been achieved. Thus for an assessment activity to be considered authentic, it is

important that consideration has been given to the relationship between the requirement to demonstrate specific individual skills and knowledge in a precise way, and the importance of producing a successful performance or product.

As, Newmann et al (1996, p. 286) assert 'authentic intellectual achievement requires construction to reach beyond retrieval and imitation of knowledge previously produced by others'. This is further supported with the notion of performance as a requirement for authentic assessment to distinguish 'between achievement that is significant and meaningful and that which is trivial and useless' (Newmann & Wehlage, 1993, p. 8).

3. Transfer of learning (skills/knowledge/attitude) required

This element seeks to determine the extent to which the skill, knowledge and attitude being assessed may have meaning beyond the confines of a single content area. As Tanner (1997, in McAlister, p. 28) states 'there should be consistency between the assessment and the real-world application for which the learner is being prepared'. Thus, in authentic work performance, knowledge may often be drawn from a range of domains, yet may be applied only within a single domain to produce successful performance. The authentic assessment activity should support the notion that knowledge and skills learnt in one area can be applied within other, often unrelated areas, to elicit successful performance. This is what Berlak (1992, p. 25), refers to as assessment relevance, 'the degree to which the assessment is related to what the learner is being prepared to do beyond the particular assessment setting'. Whilst Hattie, Biggs and Purdie (1996, p. 29) note that 'assessment should be in context and use tasks within the same domain as the target domain'. It does not preclude recognition that transfer of knowledge or skill from another domain might enhance performance.

By establishing the amount of learning transferred from the training environment to the work environment, consideration is given as to how closely the future work experience has been replicated in the assessment environment. As assessments are designed to measure the learning outcome, it is by establishing the extent of the relationship between knowledge application and its deployment that the assessment designer is able to consider the link between knowledge, skills and attitudes taught, and their application in the workplace. As Tanner (2001, p. 24) states 'assessments ought to encourage the transfer of learning', a point that is further elaborated upon by Wolf (1993, in Supovitz and Brennan, 1997, p. 472) who asserts the importance of assessment to the overall learning process: 'assessment is conceived not just as the end product, but also as an episode of the learning'. In this respect, this element also recognises that workplace performance will often deem that the production of a given performance or product is not necessarily an end in itself, but may also be one of a number of requirements needed to attain a given overall end result. In achieving this, a degree of transfer from the education environment to the workplace may be required.

4. Critical reflection and self-assessment or evaluation required

This element establishes the value and importance of both critical reflection and self-assessment or self-evaluation, as key workplace relevant performance skills. As McAlister (1994, p. 29) notes, 'monitoring their own learning through self-evaluation can enhance student learning'. Thus the on-going monitoring of our own learning via self-assessment or self-evaluation can increase overall understanding, and improve performance. Newmann and Wehlage (1993, p. 4) further note, when commenting upon their own five standards for authentic assessment, 'such criteria extend authenticity beyond simple participation in real experiences to active reflection on the meaning

beyond those experiences'. Moreover, Scott (1994, p. 5) describes self-assessment as 'a critical workplace skill'. In other words, students want to know how they are doing whilst they are performing some task. Khattri, Reeve and Kane (1998, p 14) further note that this process can be developed if 'students [are] required to explain the process by which they arrive at answers'.

Reflection enables links to be made both within and between content areas, enhancing the understanding of the processes by which satisfactory outcomes or performances are concluded. As Cizek (2000, p. 16) states 'instead of taking tests that consist solely of computation, reading or select-response formats, students are increasingly reflecting on and writing about their own mathematical, scientific and analytical thinking and problem solving abilities' This in turn, as Herrington and Oliver (2000, p. 24) note makes it possible to 'promote reflection to enable abstraction to be formed'.

The use of active critical reflection or metacognition to perform the assessment activity itself is also worth considering here. Metacognition or reflecting, is what Aschbacher (1995, in Moorcroft et al., 2000, p. 4) refers to as one of the, 'six key components identified by the National Center for Research on Evaluation, Standards, and Student Testing' and what Moorcroft et al. (2000, p.4) describe as 'the student's ability to evaluate their own progress'.

5. Accuracy in product or performance, and fidelity of assessment environment, is displayed

This element seeks to determine how central the skills and knowledge being assessed are to final work-related application. As Newmann and Wehlage (1993, p. 9) note with the first of their five standards of authentic instruction, it is possible to distinguish levels

of authenticity by asking amongst other questions, ‘to what extent are students required to use higher-order thinking skills?’ In other words, it seeks to establish the extent of intellectual input required in the development of the product or performance, as a means of determining the degree of authenticity inherent within in an activity. This provides a focus upon not only developing understanding and applying knowledge, but also the developmental process as demonstrated or evidenced by the final assessment outcome.

This element also seeks to establish whether the required outcomes are critical or peripheral to final workplace performance. An authentic assessment should simulate, and measure a real world test of ability, rather than just match items to curriculum content. In the words of Tanner (1997, p. 11):

> Authentic assessment presumes students will produce something that reflects not a narrow, compartmentalized repetition of what was presented to them, but an integrated scholarship which connects their learning housed in other disciplines and which is presented in a setting consistent with that in which the learning is likely to be most useful in the future.

This element also guides the assessment designer to consider the fidelity of the environment within which the assessment is to occur as well as the fidelity of the product or performance, as a further factor in the determination of authentic assessment design. It is noted that the nine situated learning design elements proposed by Herrington and Oliver (2000, p. 24) provide for both the provision of authentic contexts, as well as the inclusion of authentic activities. This links in with the first stage of this element, to ensure the centrality of the skills and knowledge under consideration, or as Avery (2000, p.4) notes ‘the level of task authenticity is strongly related to the level of student’s authentic intellectual performance.’

This notion is further supported by Lund (1997, p. 25), who asserts that 'assessment must involve examination of the process as well as the products of the learning'. Finally, according to Newmann and Wehlage (1993, p. 8), authentic achievement is based on three criteria, one of which is 'that students aim their work towards production of discourse, products and performances that have value or meaning beyond success in school'. This is a view shared by Wiggins (1990, p. 2) for whom, 'authentic assessments attend to whether students can craft polished, thorough and justifiable answers, performances or products'.

6. Fidelity of assessment tools used

The tools being used to conduct the assessment should replicate those used in the real world environment. For Berlak (1992, 2001, p. 24) 'the hallmark of authentic assessment practices is their harmony with real world circumstances'. This then includes the use of any tools that would be considered to be appropriate to the work environment. In large part, if authenticity in assessment is a function of the nature of task to be undertaken and the environment within which it is to occur, it is vital to ensure also that any tools used to undertake it are also authentic to the task. As Newmann, Lopez and Bryk (1998, p.3) note 'students produce more authentic work when given challenging, engaging tasks, particularly those that have real world connections'.

This point is further elaborated by McLellan (1994, p. 6), who asserts that 'if the assessment occurs within the context for which it is intended to be used, then such a context is sufficient as it is usually either a replica of the appropriate environment, or a

contextual anchor which reflects the conventions of the environment'. In this respect, the authenticity of the tools applied is maximised.

It is noteworthy that, depending upon the circumstances and nature of the assessment, the definition of tools may include broader cultural elements such as language. Therefore, according to Northcote and Kendle (2000, p.5), 'to enable a task to be as authentic as possible, culturally appropriate language, graphics and topics are used to make the students feel more familiar with the assessment task'.

7. Discussion and feedback required

The ability to discuss, give and receive feedback is also critical to workplace performance. Newmann and Wehlage (1993, p. 4), state that the levels of authenticity in an assessment should include 'the extents to which students are expected to discuss, learn, and understand the substance of the subject'. In addition, students must also, as McAlister (1994, p. 29) notes 'be continually monitoring their own learning through self-evaluation [which] can enhance student learning'.

Thus the value of feedback, not just guidance, as a means of enhancing performance and seeking out external sources for gathering critical data and determining the areas where improvement can be made, is vital to improved performance. To this end, as Northcote and Kendle (2000, p.8) assert 'it is extremely helpful to build in opportunities for feedback in assessment'. Herrington and Oliver (2000, p.4) also note the role and value of articulation to enable tacit knowledge to be made explicit.

The value of being able to seek out feedback and engage in discussion upon an activity is considered to have meaning and value beyond the successful completion of an

assessment activity. In this respect, as a critical skill in successful workplace performance, it must be viewed as one of the critical elements in the determination of an authentic performance. As Newman and Wehlage (1993, p. 4), further note, it is one means of ensuring that the assessment activity may have 'value and meaning beyond the classroom'. Thus, this element seeks to, as Newman and Wehlage (1993, p. 4) further assert 'extend authenticity beyond simple participation in 'real' experiences to active reflection on the meaning beyond those experiences'.

It should be noted that whilst reflection is dealt with specifically within the fourth critical element, it is acknowledged that there will always be a requirement for a degree of reflection to have occurred to enable appropriate discussion to take place and feedback to occur.

8. Collaboration required

The ability to collaborate is also important in workplace performance. The value of collaboration, as a means of enhancing performance and seeking out external sources for gathering critical data and determining the areas where improvement can be made, is vital to improved performance. As Northcote and Kendle (2000, p.6) state 'the socio-cognitive value of collaborative learning is one that is becoming increasingly recognised and also offers students access to multiple points of view as well as some useful opportunities for modelling'.

Lebow and Wager (1994, p. 239) advocate collaboration as one response to the requirement to utilise a 'holistic and generative approach to education and the use of technology to assist students in developing higher order thinking skills and important

long term dispositions to learning'. As they further state, collaboration provides students with the opportunities to engage in activities that,

> …a) shift from all students learning the same things to different students learning different things; b) create group problem-solving situations that give students responsibility for contributing to each other's learning; and c) help students see the value of what they are learning and choose to share.
> (Lebow & Wager, 1994, p.241)

Herrington and Oliver (2000, p.4) support this view when they include 'the support [of] collaborative construction of knowledge' within the nine situated learning design elements. Collaboration then, much as with the ability to give and receive feedback, is as Reeves, Herrington and Oliver (2002, p. 564) describe 'integral to the task, both within the course and the real world, rather than achievable by an individual learner'. More particularly, the move to the greater integration of web technology within a blended approach to learning engagement means that as Reeves, Herrington and Oliver (2002, p. 566) note 'it is possible to include more engaging collaborative activity'.

Thus the original eight critical elements are devised and have now been considered in detail. The next stage is that of the review of these critical elements by acknowledged experts in the field of instructional and educational design.

Expert review

Expert review of the critical elements shown in Figure 4.2 was sought from three practitioners currently employed in this field. The process for obtaining this feedback, as previously described in Chapter 3, was mainly by means of synchronous verbal discussion, but some occurred via e-mail contact.

These experts provided extensive feedback on each of the elements selected on both their individual value as well as the value that they might hold as a part of a framework. On completion of this collaborative process, the feedback received was collated and analysed and the framework of draft elements was further revised in the light of that feedback. The responses of each of the experts were weighted equally.

Set out below, is an analysis of the feedback received from each of the expert individual reviewers. This feedback is presented in the form of an introduction of the overall views and opinions of the reviewer upon the critical elements, followed by a detailed outline of their feedback against each of the eight elements. Finally, a conclusion is provided that summarises the individual responses of the three expert reviewers into an overall collated response against each of the eight individual elements.

Expert Reviewer 1

Expert Reviewer 1 is a university lecturer in information and communication technology and multimedia with a particular interest in on-line course design and generic skill development. His particular research interests include authentic learning, student-centred learning and self and peer assessment.

Whilst Expert Reviewer 1 agreed overall with the critical elements, he felt that they could have benefited from the addition of a ninth element such as *client*. In his opinion, the role of the client in terms of the entity for whom a final outcome is produced should be considered as an important element in the determination of the authenticity of the assessment experience.

He also stated that he had expected to see the emphasis placed upon nouns instead of verbs. To this end he suggested that the nouns within each of the elements should be underlined and descriptors be used, in order to more clearly focus attention on that which was critical in each element, that is, *challenge, performance, transfer, critical reflection & self assessment, accuracy, fidelity, discussion and collaboration.*

1. Degree of challenge(s) presented to the assessed student

Expert Reviewer 1 agreed that this was a critical element but thought that the client would also have issues that will exist in the 'real world' situation. He also considered that the degree of challenge was a job in itself such as having to analyse the real problem. Finally, he noted that consideration should be given to combining this element with accuracy in product or performance.

2. Performance, or product, as final assessment outcome

Expert Reviewer 1 agreed with this element. He further commented that 'the course of study that they undertake uses the product itself that the trainees produce as well as the process that they apply as the means of marking the assessment.' Furthermore, he also felt that this should also be reflected in the marking process, as shown below:

'Process→Critical Element→Authentic Assessment Task→Assessment Outcome'

3. Transfer of learning (skills/knowledge/attitude) required

Expert Reviewer 1 stressed that this was a very important element in authenticity in assessment. He argued that in authentic assessment students should not be learning skills in an abstract manner, but they should be directly applying learnt skills. He also argued that transfer of learning related more to an outcome or output of learning.

4. Critical reflection & self-assessment or evaluation required

Expert Reviewer 1 agreed strongly with the inclusion of this element as he felt that it encouraged the development of reflective practitioners. He believed very strongly in the role of reflection in assessment. He stated that within his own teaching environment he expected his students to reflect on a weekly basis on their practice. Furthermore he also indicated that had constructed on-line tools to both enable and assist this process. From his perspective as a teacher, he noted that if he did not have the ability to review the outcome of weekly student reflection, then he would have no means to follow a student's progress.

5. Accuracy in product or performance, and fidelity of assessment environment, is displayed

Whilst Expert Reviewer 1 supported the inclusion of this element, he felt that consideration should be given to combining it with Degree of Challenge.

6. Fidelity of assessment tools used

Expert Reviewer 1 believed that this element should be re-titled 'fidelity of tools used to conduct assessment' and that this would make it easier to comprehend. He also believed that this element should be flexible enough to reflect any relevant industry standards. In this regard, he considered that students should undertake what he considered to be 'industry-strength' assessment items.

7. Discussion and feedback required

Expert Reviewer 1 agreed strongly with the inclusion of this element. He thought that students should be encouraged to be aware of a range of generic skills, of which the ability to discuss and provide feedback was just one. Other generic skills he identified

as important elements were body-language, communication skill and cultural awareness. Expert Reviewer 1 noted that consideration should be given to the combination of this element with collaboration.

8. Collaboration required

Finally, Expert Reviewer 1 again agreed strongly for the inclusion of this element. He believed that collaboration was in itself a valid indicator of authentic activity and should be seen to encompass problem-solving ability. As noted above, he expressed a preference for combining this element with that of discussion and feedback required.

Expert Reviewer 2

Expert Reviewer 2 is also employed as a university lecturer in the fields of assessment and educational technology and has experience as an instructional designer. Her particular research interests include online unit design, online assessment and staff development.

Overall, Expert Reviewer 2 considered that each of the elements identified were critical. In addition, she felt that the crucial elements, as applied at her own institution, were collaboration, challenge and feedback. Expert Reviewer 2 also agreed with the use of the word *elements* as opposed to *components.* However, whilst she did note that that there was a degree of overlap within some elements listed, she felt that each made sense in its own right, and that each built upon the work of others.

In addition, Expert Reviewer 2 felt that it might be possible, from these elements, to develop criteria for on-line assessment for use within the TAFE system. She felt that such a tool might have a useful application in the development of activities for the

assessment of vocational outcomes, particularly resulting from the TAFE systems increased use of on-line content delivery.

1. Degree of challenge(s) presented to the assessed student

Expert Reviewer 2 supported the inclusion of this element and felt that its description was comprehensive. However, she also felt that it did not link directly enough to the issue of challenge as suggested in the title. It was felt that at present the link to challenge was more implicit and that it should be more explicitly stated. As a result the descriptor should be amended to reflect this.

2. Performance or product as final assessment outcome

Expert Reviewer 2 concurred with the inclusion of this element and noted that it further supported authenticity with the implication that final performance or product was of value.

3. Transfer of learning (skills/knowledge/attitude) required

Expert Reviewer 2 noted the link between this element and performance or product as the final assessment outcome. However, she considered that it was a difficult outcome to assess. She commented that once knowledge had been learnt, it was often difficult to follow-up and track in the workplace. Finally, she noted that there was a role for subsequent validation of this element.

4. Critical reflection & self-assessment or evaluation required

Reviewer 2 noted that this was an element that was often neglected, as it was assumed that it was being undertaken by educators/teachers or a facilitator directly with students.

She also stated that this element should include metacognition. Expert Reviewer 2 felt that all self-assessment leads ultimately to the process of metacognition.

5. Accuracy in product or performance, and fidelity of assessment environment, is displayed

In addition, Expert Reviewer 2 supported the inclusion of this element, but outlined that in her opinion the title contained two distinct ideas, and that accuracy in product or performance should be considered separately to the fidelity of the assessment environment.

6. Fidelity of assessment tools used

This element received support from Expert Reviewer 2. In particular, she felt that it was important to make mention of the tools to be used within the assessment environment, and the importance of ensuring that they replicated those to be used in the workplace. She also noted that if the tools are simulated such as using multimedia that they could be considered to be as effective as the real tools. Finally, she felt that the descriptor should be amended to include the use of simulated tools.

7. Discussion and feedback required

Expert Reviewer 2 supported the inclusion of this element. She believed that whilst teachers in the school environment conducted numerous summative assessment activities, they paid less attention to formative assessments which were often more valuable as a means of providing feedback to the educator. She also felt that it was important to stress the requirement to provide on-going feedback, and that both discussion and feedback were of vital importance to aid the process of self-reflection and metacognition. However, she also considered that some teachers provided too much

feedback or guidance about assessments and that this only served to make students become too dependent. Finally, she noted that as well as providing too much feedback, the criteria on marking rubrics can be too specific and, thereby, not allow sufficient scope for the use of the student's imagination.

8. Collaboration required

Finally, Expert Reviewer 2 stated that collaboration was 'absolutely essential'. In her opinion, few authentic environments exist without the need for collaboration. She also outlined the requirement for two types of collaboration, namely, 'working with one group of people', and, 'obtaining information from another group of people.'

Expert Reviewer 3

Expert Reviewer 3 is a senior instructional designer with significant experience in the design and development of authentic learning outcomes particularly for vocational training. In addition, she has a high level of experience in the design and development of authentic assessment activities to determine that these learning outcomes have been achieved.

Expert Reviewer 3 began by asserting that, in her opinion, the critical elements were appropriate.

1. Degree of challenge(s) presented to the assessed student

Expert Reviewer 3 agreed with the inclusion of this element and commented that 'in the real world; a prime regulator of authenticity would be workplace relevance'. In addition, she felt that 'the degree of challenge implicit within an assessment, should seek to ensure that a student was able to both appreciate a situation and make informed

decisions based on that appreciation, within the often simulated assessment environment'.

2. Performance or product as final assessment outcome

This element was supported on the basis that it is the 'production of performances or products, in the workplace that are often the principal means of demonstrating both capability and competence'.

3. Transfer of learning (skills/knowledge/attitude) required

Expert Reviewer 3 supported the inclusion of this element, and noted that it was important to acknowledge the value of skills demonstrated, above pure knowledge regurgitation. In addition, she affirmed that employers were less interested in 'essay content' than they were in the skills required to write an essay, such as time-management and self-discipline.

4. Critical reflection & self-assessment or evaluation required

Expert Reviewer 3 strongly supported the inclusion of this element on the basis of the fundamental importance of both critical reflection and self-assessment in the consideration of authenticity.

5. Accuracy in product or performance, and fidelity of assessment environment, is displayed

The inclusion of this element was supported. However, she felt that it may be considered as too similar to the first element, the degree of challenge(s) presented to the assessed student.

6. Fidelity of assessment tools used

Expert Reviewer 3 supported the inclusion of this element, and stated that she considered the fidelity of the assessment tool as a key determinant in defining authenticity in assessment.

7. Discussion and feedback required

This element received the support of Expert 3. In addition, she noted that the richness of feedback should be used as a step in the learning process rather than just as a means providing right/wrong feedback. She further stated that 'rich feedback which refreshes an answer is a step in the learning process other than an end-point'; in this respect she considered that the process of giving and obtaining feedback was an important formative component of the authentic assessment.

8. Collaboration required

Finally, Expert Reviewer 3 agreed with the inclusion of an element on collaboration. She noted that, even in a simulated environment such as an assessment environment, students needed to appreciate the situation, then, based on that appreciation; demonstrate that they are able to make informed decisions.

Summary of feedback from expert reviewers

Whilst all three experts consulted expressed support for the critical elements listed, Expert Reviewer 2 made particular reference to the use of the word *element* as an appropriate descriptor. Expert Reviewer 1, in particular, also felt that the elements could have benefited from the addition of a ninth element such as *client*. He felt that it was important that the role of the *client* in replicating authentic performance should be

included. Further to this he suggested that the nouns in the descriptors should be underlined to focus attention on to what he considered to be critical within each element such as *challenge, performance, transfer, critical reflection* and *self assessment, accuracy, fidelity, discussion and collaboration*. This was supported in part by Expert Reviewer 2, where she noted that at her institution the crucial elements applied were *collaboration*, *challenge* and *feedback*, two of which are included, three, if discussion were to be replaced with feedback. Finally, it was also noted by all the expert reviewers that there was some overlap in a number of the elements listed.

1. Degree of challenge(s) presented to the assessed student

All three experts consulted found agreement with the inclusion of this element. In particular, support was given to the idea that the degree of challenge is a task in its own right, in terms of a student having to make an analysis of a real problem. This was supported by the notion that in a real world workplace, workplace relevance would be considered a prime regulator of authenticity. In addition, the degree of challenge implicit within an assessment, should seek to ensure that a student is able to both appreciate a situation and make informed decisions based on that appreciation. It was noted, however, that this element might not have linked directly enough to the issue of challenge as suggested in the title. In this respect, it was felt the link to challenge was too implicit and should be more explicit. It was suggested that the descriptor be amended to reflect this. Some limited discussion also occurred on the notion of combining this element with accuracy in product or performance.

2. Performance, or product, as final assessment outcome

All expert reviewers also generally supported this element, although, Expert Reviewer 1 introduced the notion that the assessment should use the product or performance that the trainees produce, as well as the process that they apply, as the means of making the assessment. This would also be reflected in the marking process, as shown below:

'Process→Critical Element→Authentic Assessment Task→Assessment Outcome'

It was also noted that, it is the production of performances or products that are the principal means of demonstrating both capability and competence.

3. Transfer of learning (skills/knowledge/attitude) required

Again, this element received support for its inclusion from all three expert reviewers. It was generally considered by the Expert Reviewers that to be considered authentic an assessment should require that students apply learnt skills directly related to a relevant task. In this regard acknowledgment was made of the importance of recognising the value of skills demonstrated, above pure knowledge regurgitation. In addition, it was also noted that employers would be less likely to be interested in the content of an essay than the skills required to write the essay. In this respect, it appears that the transfer of learning relates more to an outcome or output of learning.

Finally, a link with performance or product as the final assessment outcome was acknowledged, and it was noted that once learnt, knowledge retention is often difficult to follow-up and track in the workplace.

4. Critical reflection and self-assessment or evaluation required

This element was strongly supported by all three expert reviewers, and it was felt that its inclusion would encourage the development of reflective practitioners. It was noted that teachers needed the outcome of regular student reflection, as a means of assessing a student's progress. It was considered that this element was often neglected, in the often mistaken assumption that it occurred naturally. Finally, it was considered that this element should include reference to metacognition.

5. Accuracy in product or performance, and fidelity of assessment environment, is displayed

All three expert reviewers supported the inclusion of this element, noting that both *accuracy in product or performance* and the need to *display fidelity of assessment environment* were critical links to an authentic workplace performance. It was agreed that often within the workplace it is the degree to which a final product or performance is in accord with that expected that is the overall determinant of its success. In this regard, the method employed to achieve the performance, as long as it fell within reasonable boundaries, was of much less final value to the workplace than the product that would result from it. In the same context, it was agreed that if the assessment activity was going to authentically determine the ability of an individual or team to produce or perform, then the environment within which that *product* or *performance* was realised would by necessity have to be authentic, that is have a high degree of fidelity with the real to be valid.

It was also felt by the Expert Reviewers that to some extent that this element could be could be combined with that of *degree of challenge*. Attention was also drawn to the

fact that the title contained two distinct ideas, and that accuracy in product or performance should be considered separately to fidelity of assessment environment.

6. Fidelity of assessment tools used

Each expert reviewer expressed support for this element, although, Expert Reviewer 1 suggested that it be re-titled as *fidelity of tools* used to conduct assessment. It was also noted that this element needed to be able to reflect any appropriate industry standards that may apply. However, the importance of ensuring that the tools used in the assessment replicated those used in the workplace received general support. Finally, it was stated that simulated tools for example using multimedia should be considered to be as effective as the real tools, in this respect it was suggested that the descriptor be amended to include the use of simulated tools.

7. Discussion and feedback required

All three expert reviewers supported the inclusion of this element. It was stated that students should be encouraged to be aware of a range of generic skills, including the ability to discuss and provide feedback, as well as body-language, communication skill and cultural awareness. It was also noted that, whilst many educational courses used numerous summative assessment activities, less attention was often given to the value of formative assessments as a means of providing feedback to the educator. The role and richness of feedback was also stressed, in particular the importance of the requirement to provide on-going feedback, as being of vital importance, to aid the process of self-reflection and metacognition. The richness of feedback was also suggested as a step in the learning process rather than just as a means of providing right/wrong feedback.

Finally, it was requested that consideration be given to combining this element with collaboration.

8. Collaboration required

Each expert reviewer supported the inclusion of this element. *Collaboration* was viewed as a concept that was not just valuable to assessment performance in the learning environment, but crucial a broader more generic skill that would need to be consistently applied in the work environment. Few authentic environments exist without a requirement for collaboration, the requirement to work with other people on complex tasks.

Even in a simulated environment such as an assessment environment, students needed to be able to make an appreciation of a situation, and then demonstrate that they were able to make informed decisions. Finally, whilst it was considered that this element could be combined with *discussion and feedback required* it was decided that sufficient difference existed between the two to justify their continued separation into the two critical elements.

Revision of critical elements from expert review

Table 4.3 summarises the revised framework of critical elements incorporating the expert review. This table is divided into three columns, the first of which, *Critical Element*, sets out the original eight critical elements as reviewed by the experts. The second column, *Expert Reviewer Feedback*, provides a summary of the collated expert responses concerning that particular element. The final column, *Elements expressed as a critical question,* sets out the critical element in their post expert review format. The

terms are those which are used in the design and development phase and which form the basis for the final evaluation of the elements as determinants of authentic practice in assessment design.

It should be noted because of the iterative nature of a design based research study, it was determined that whilst all expert feedback was acknowledged at this stage, analysis of feedback meant that on balance not all expert feedback could not be used to amend particular elements at this stage. However, this feedback can be retained for further review and discussion in subsequent studies.

Table 4.3: Revision of critical elements from expert reviewer feedback to produce the critical questions

	Critical elements	***Expert reviewer feedback***		***Elements expressed as critical question***
1.	Degree of challenge(s) presented to the assessed student.	The degree of challenge implicit within an assessment should seek to ensure both an appreciation of a situation and, ensure that a student made informed decisions based on the appreciation. The link to 'challenge' should be more explicit, and the descriptor needs to be amended to reflect this.	1.	To what extent does the assessment activity *challenge* the assessed student?
2.	Performance, or product, as final assessment outcome.	The assessment should use the product or performance, as well as the process applied, to make the assessment.' This should be reflected in the marking process, as shown below: 'Process→Critical Element→Authentic Assessment Task→Assessment Outcome' Production of performances or products, that is the principle means of demonstrating both capability and competence.	2.	Is a *performance, or product*, required as a final assessment outcome?
3.	Transfer of learning (skills/knowledge/attitude) required.	Acknowledge the value of 'skills demonstrated', above 'pure knowledge regurgitation'. Link with 'Performance or product as the final assessment outcome'.	3.	Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?

	Critical elements	***Expert reviewer feedback***		***Elements expressed as critical question***
4.	Critical reflection & self-assessment or evaluation required.	Include reference to 'metacognition'.	4.	Does the assessment activity require that *metacognition,* is demonstrated by means of critical reflection, self-assessment or evaluation?
5.	Accuracy in product or performance, and fidelity of assessment environment, is displayed.	Consider combining with 'Degree of Challenge'. Note that the title contains two distinct ideas, such as 'Accuracy in product or performance' and 'fidelity of assessment environment'. Note concern that this element too similar to — 'Degree of challenge(s) presented to the assessed student'.	5.	Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?
6.	Fidelity of assessment tools used.	Consider re-titling as 'Fidelity of tools used to conduct assessment', and ability to reflect an appropriate industry standards that may apply. Ensure that the tools used in the assessment replicate those used in the workplace. Descriptor should be amended to include the use of simulated tools.	6.	Is fidelity required in the assessment environment? And the assessment tools (actual or simulated)?
7.	Discussion and feedback required.	Note 'richness of feedback', as a step in the learning process rather than just as a means providing 'right/wrong' feedback. Consider combining this element with 'Collaboration'.	7.	Does the assessment activity require *discussion* and *feedback*?
8.	Collaboration required.	'Collaboration' a good phrase encompasses 'problem-solving' ability. Authentic environments exist with a requirement for collaboration, for example The requirement to 'work with other people' and, 'obtain information from other people.' Even in a 'simulated environment' such as an assessment environment, students need to be able to make an appreciation of a situation, and demonstrate ability to make 'informed decisions'. Consider combining this element with 'Discussion and feedback required.'	8.	Does the assessment activity require that students *collaborate*?

From critical elements to critical questions — A summary

As Table 4.3 encapsulates, what emerged from the process of expert review in Phase 2.2 was a change in format. From being a series of eight statements, the critical elements became instead, a series of eight critical questions that an educational designer could more easily apply to the design of both authentic educational content as well as assessment tools.

In addition, it was possible, largely based upon the responses of the expert reviewers, to highlight key words within each of the revised elements. This provided an additional and simplified indication as to that aspect of each of the elements that may be considered to be most crucial. These revised elements, in the format described, are set out below:

The critical questions

1. To what extent does the assessment activity *challenge* the assessed student?

2. Is a *performance, or product*, required as a final assessment outcome?

3. Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?

4. Does the assessment activity require that *metacognition*, is demonstrated, by means of critical reflection, self-assessment or evaluation?

5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?

6. Is *fidelity* required in the assessment environment? And the assessment tools (actual or simulated)?

7. Does the assessment activity require *discussion* and *feedback*?

8. Does the assessment activity require that students *collaborate*?

The next chapter, Chapter 5, considers Phase 2.3 and describes how this framework can be applied to the design of a discrete module of learning, namely, 'Evaluating Educational Multimedia' of the Australian Army's Educational Multimedia Developers Course. The chapter will further describe the purpose of this course, and consider the role of Module 10 within this course. The subsequent focus is on both the design and development of the course content and assessment tasks through the application of the framework of critical questions presented in Figure 4.3.

CHAPTER 5: APPLYING THE CRITICAL QUESTIONS OF AUTHENTIC ASSESSMENT IN THE DESIGN OF A LEARNING MODULE

Development of — *Evaluating Educational Multimedia*

Introduction

In Chapter 4, an initial list of the critical elements of authentic assessment was derived from the literature. By means of a process of research, review and expert feedback these were further refined in an iterative process, and became the critical questions that an educational designer would need to consider in the design and development of a learning and assessment activity.

This chapter seeks to address the first of the subordinate research questions: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex authentic tasks*, and describe the ways in which these elements, and the questions derived from them, were used in the design of a specific module of learning — Evaluating Educational Multimedia, a module of the Australian Army Training Technology Centre's Computer Based Learning Practitioners course (Appendix 4).

The intent of this module was that of providing training to newly appointed Army instructional designers in the general principles and practices of the evaluation of educational multimedia and its particular application to the evaluation of Army computer based learning packages. The module itself was comprised of three distinct learning outcomes (LO) with their related assessment criteria (AC) as set out below:

- LO 1.0 Explain educational multimedia evaluation models
  - AC 1.1 Explain educational multimedia
  - AC 1.2 Outline the main approaches in the delivery of educational multimedia
  - AC 1.3 Describe the types of evidence to be gathered from the evaluation of educational multimedia
- LO 2.0 Outline the structure of an educational multimedia report
  - AC 2.1 List the approaches and methods for evaluating elements of an educational multimedia report
  - AC 2.2 Identify the elements to be evaluated
  - AC 2.3 Describe the components of a revision plan
- LO 3.0 Apply the process of educational multimedia evaluation to a Training Technology Centre developed computer based learning product
  - AC 3.1 Summative assessment — Identify the aims of a particular computer based learning package
  - AC 3.2 Report on the value of that computer based learning package

It was upon these assessment criteria that the assessment for this module was based.

The re-design of *Evaluating Educational Multimedia*

As discussed in Chapter 3, this module was developed as the re-design and re-development of an existing learning module. On completion of a Defence provided training course, the students are normally assessed at the training centre prior to their departure, to obtain information about their perspective on the course delivery

experience that they have just received. They are also assessed again, usually within the period three to six months after completing the course, in their workplace, to establish how well they feel that the training that they have received has equipped them for their current employment. At the same time, their workplace supervisor is also interviewed to provide information as to how well the course that they have just undertaken has prepared them for the workplace in which they are employed. These processes are established within the Evaluation Phase of the Defence Training System (ADFP, 7.0.2, 2007). It was the evidence gained from the students who had undertaken the previous version of this course, including this module, that acted as the catalyst for the course and module's re-design.

The previous students had made a number of comments with regard to the teaching of this course, and these included:

- The facilities in the training classroom were not adequate or workplace relevant, and were not equipped with the tools to be used in the workplace.
- Students had no access to instructors out of training hours.
- The students were not provided with the learning content to be covered ahead of time.
- The face to face instructional techniques employed by staff could be improved.
- There were numerous instances of repetition throughout the course.
- The students thought that more time could have been spent on applying the instructional content and less time just learning theory.
- The students thought that better use could have been made of multimedia developed resources to support the print-based modules.

- As the assessment items were all theory-based, there had been duplication of assessment items, which were generally ‘regurgitation’ exercises.

On the basis of that feedback the following recommendations for improvements to the course were made:

- Investigate upgrading the facilities of the training classroom or use a different learning environment, to be more consistent with the professional working environment and its tools.
- Ensure teachers are experienced in teaching in an adult learning environment.
- Investigate the nature of the content and delivery of the assessment items to make them more relevant to workplace performance.
- If practical, provide students with learning materials in advance.
- Investigate changing the teaching delivery method for the course.

Table 5.1 (below) provides a description as to how these comments, allied to the critical questions, became the basis for the redesign of this module. In particular, and against each of the critical questions, it describes the assessment methodology applied in the module’s previous iteration, the proposal for its redesign against the critical question and the rationale as to the designer’s intent in making the described changes.

Table 5.1: Proposed application of the critical questions to the re-design of Module 10

Critical Question	*Assessment before*	*Proposal for assessment after re-design*	*How and why changed — designer rationale*
1. To what extent does the assessment activity *challenge* the assessed student?	Assessment requires the student to answer non-applied theory questions as a test of memory.	Assessment to require students to apply the theoretical content in practice.	The intention was to increase the degree of challenge on the student by expecting them to apply what they had learnt in theory to achieve an applicable outcome.
2. Is a *performance, or product*, required as a final assessment outcome?	Assessment requires student to answer questions in narrative (sentences/paragraphs) format.	Assessment to require student to design and develop a workplace applicable tool.	The intention was to ensure a stronger link between knowing theory and applying it to the design and development of a workplace applied tool.
3. Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?	Assessment requires limited transfer of knowledge in undertaking of a non-applied theory test.	Assessment to require transfer of theoretical knowledge in the design and development of a workplace applicable tool.	The intention was to reinforce the transfer of theoretical knowledge with its application to the design and development of a workplace applicable tool.
4. Does the assessment activity require that *metacognition*, is demonstrated, by means of critical reflection, self-assessment or evaluation?	Assessment has limited or no requirement for metacognition.	Assessment to require that student reflects critically and self-assesses their own designed outcome.	The intention was to give the student an opportunity to be able to reflect on the design decisions that they had made and self-assess the outcome in the context of both theory and the work of colleagues.
5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?	Outcome of assessment is determined by a teacher, who is effectively functioning in the role of client but this is not overtly obvious to the students.	Student attention to be drawn to the fact that the success of the assessment outcome will be determined by its application in practice and that the teacher, functioning as a client will assess it on this basis.	In the original assessment activity the students were not made aware of the function of the teacher as a client and instead the teacher was viewed more in the traditional role of assessor. In the revised assessment student attention was drawn to the fact that the assessment tool would be applied in practice and assessed by the teacher in that context.

Critical Question	*Assessment before*	*Proposal for assessment after re-design*	*How and why changed — designer rationale*
6. Is *fidelity* required in the assessment environment? And the assessment tools (actual or simulated)?	Assessment makes limited or no attempt to situate the activity in a workplace relevant context and does not require application of actual workplace tools (software).	Assessment activity to be situated within a high fidelity working environment and the tools applied in practice (software) to be those applied in the workplace.	The intention of the revised assessment is to ensure that it is conducted in an environment that is as close as possible to the actual workplace environment. In addition, the software tools made available to students are to be the same as those used in the workplace.
7. Does the assessment activity require *discussion* and *feedback*?	Assessment provides little or no opportunity for either *discussion* or *feedback.*	The requirement for *discussion* and *feedback* is integral to successful assessment performance.	The revised assessment activity to be re-modelled to ensure that students have to *discuss* and receive *feedback* from students and the teacher.
8. Does the assessment activity require that students *collaborate*?	Assessment provides little or no opportunity for *collaboration.*	The opportunity for student's *collaboration* is integral to successful assessment performance.	The revision to the assessment activity to ensure that students are expected to *collaborate* with one another in the completion of a successful assessment performance.

The following section describes in more detail, on a critical question by question basis, the ways in which the overall design of the module would be revised to be consistent with these eight critical questions, and the student feedback received on the evaluation of previous course and module delivery.

1. To what extent does the assessment activity *challenge* the assessed student?

Whilst it is acknowledged that challenge can be manifest in many ways, for the purpose of this particular module, challenge is measured in terms of the degree of difficulty inherent in achieving the final outcome, in particular the degree to which the student has to be able to demonstrate the synthesis of theory with the skills and knowledge that they have acquired. The feedback received from students indicated a strong concern that the

design of the assessment for the previous module had required the student to answer non-applied theory questions as a test of memory.

As this element requires both a considered analysis of the task, as well as the subsequent selection of the appropriate response, it was intended that the redesign of this element would reflect the authenticity of real world situations and tasks where the challenge of providing a successful outcome will be dependent upon a range of factors. It may be considered to represent the degree of difficulty to be faced by a student in developing a successful outcome. To this end the assessment should require students to apply the theoretical content in practice.

In this regard, the successful completion of the module was ultimately judged with an assessment activity that required the completion of a product that was directly applicable in a work context and had a demonstrable value. The degree of challenge for the purposes of this particular activity was considered as being high, particularly in the light of the fact that for the students, this would be the first time that they would have given formal consideration to the evaluation of such a piece of educational content, and then had to produce a tool for its evaluation.

2. Is a *performance, or product*, required as a final assessment outcome?

As stated within the description of the critical questions, under this particular element a determination is made as to the extent to which the assessment activity requires a performance or product as its outcome. Often, within a work environment, the application of skills and knowledge is judged by means of a completed performance or product. In some respects, the requirement to demonstrate the particular ways in which specific skills and knowledge are applied to the achievement of that outcome may be

secondary to the requirement to produce the outcome. For an assessment activity to be considered authentic, some manifest outcome, in performance or product terms should be apparent.

The student feedback from the previously delivered version of this module informed us that the assessment had required student to answer questions in narrative (sentences/paragraphs) format.

Within the context of this particular critical question, and for the purposes of this module, the revised student's final assessment would require that the students design and develop a workplace applicable tool, and that the outcome of the assessment would be based upon a review by the teacher, of that final product, that is the evaluation tool that the students produced. Both the evaluation tools itself, and how well that tool could be used in the evaluation of a piece of educational multimedia, were the means by which the assessment was judged.

3. Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?

Whilst, this criterion seeks to determine the extent to which the skill, knowledge or attitude being assessed may have meaning beyond the curriculum area. It recognises that in authentic performance, these will often be drawn from a range of discreet domains that may need to be applied within a single area or domain to elicit a successful performance. The authentic assessment activity should demonstrate the transfer of any theoretical knowledge or foundation skills with the outcome of a successful product or performance.

The original design of the learning module's assessment had required only the limited transfer of knowledge in the undertaking of a non-applied theory test.

As it is by determining the amount of learning transfer that occurs from the training environment to the work environment that consideration is given as to how closely the workplace experience can be replicated in the assessment environment, then the redesign of this assessment would need to allow for that transfer of theoretical knowledge to the design and development of a workplace applicable tool.

As assessments are designed to measure the training outcome it is by establishing the extent of the relationship between knowledge application and its deployment, that the assessment designer is able to consider the link between knowledge, skills and attitudes taught, and their application in the workplace.

Thus within the context of the assessment activity for Module 10, again, students were expected to produce a specific and applicable tool that could be directly utilised within a work environment. However, the successful design and development of such a tool was dependent upon the application of the knowledge acquired during the completion of the learning module.

4. Does the assessment activity require that *metacognition* is demonstrated, by means of critical reflection, self-assessment or evaluation?

In general terms, the on-going monitoring of our own learning via self-assessment or self-evaluation can increase overall understanding, and improve performance, a critical component of the improvement in outcome for this course overall and Module 10 in particular. Reflection also enables links to be made within and between content areas,

thus enhancing the understanding of the processes by which satisfactory outcomes or performances are concluded. In this way, consideration is given to the use of active critical reflection to perform the role of the assessment activity itself.

Feedback received from students at evaluation suggested that the design of the assessment activity for the original module had provided little or no opportunity for metacognition. It was decided that the revised assessment would require that students reflects critically and self-assesses their own designed outcome.

Within the redesigned Module 10 students were provided with a number of opportunities to both reflect upon their own work, as well as offer feedback to colleagues, and further to reflect upon the feedback of their colleagues upon the work that they had produced. In particular, the opportunity to reflect critically upon their own work was available at the completion of part one of the assessment activity, where, once they had applied the first version of their evaluation tool, they could critically reflect upon it prior to seeking the feedback of others, as illustrated in Figure 5.1.

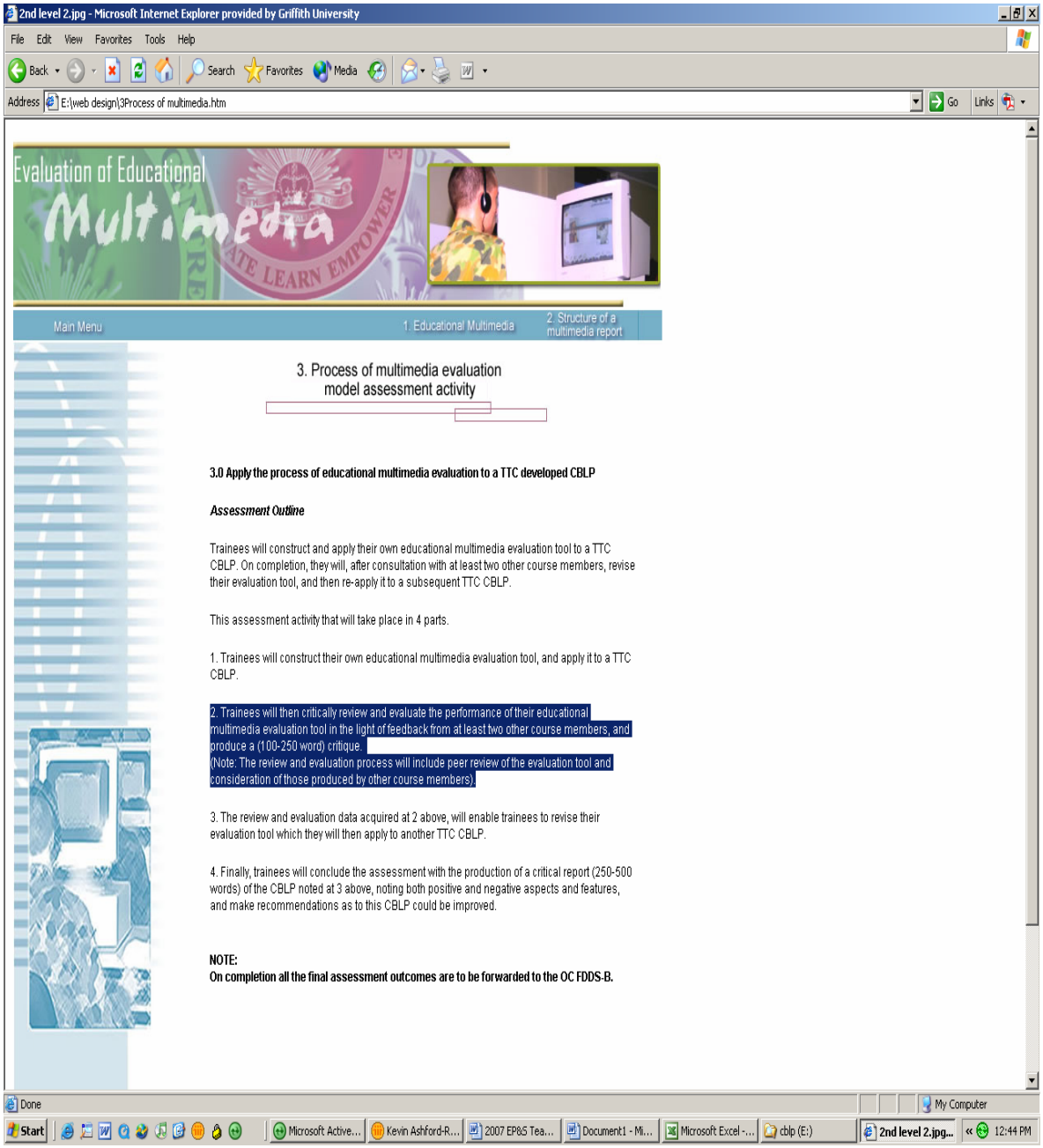


Figure 5.1: Apply the process of educational multimedia evaluation to the Army's Training Technology Centre developed Computer Based Learning Practitioners Course

5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?

Students stated that the outcome of the original assessment activity had been determined by a teacher, who whilst in effect functioning in the role of client, this was not made

clear to the students. For the purposes of the redesign of the assessment activity, students were reminded that the final outcome would be assessed by the teacher/researcher, and that the accuracy of the tool or product that they developed, in terms of its ability to undertake the required evaluation, would be critical to their success in completing the module. Student's attention then was drawn to the fact that the success of the assessment outcome was determined by its application in practice and that the teacher, functioning as a client, would be assessing it on this basis.

This is consistent with the requirement of an authentic assessment to simulate, and measure, a real world test of ability, as opposed to just matching items to curriculum content, as set out in Figures 5.2. In this respect the term *client* is used to represent the person, persons or organisation who will ultimately be the recipient of the final product or performance. Whilst it is acknowledged that the use of the term *client* might imply a commercial transaction, it is taken as having a broader meaning here as the term which most closely indicates the relationship between the task to be undertaken and the fact that the task will be reviewed, and when appropriate, accepted by another. Whilst other terms were considered here, namely, *constituent* or *supervisor*, it was felt that neither of these gave sufficient emphasis to the broad range of arenas from which a workplace client might come, thus, whilst acknowledging that the term might not in all cases be quite specific enough, it was accepted as the closest available descriptor to determine the relationship between the assessed and the assessor in the authentic assessment environment.

The second part of this criterion determines the value and fidelity of the environment within which the assessment activity is conducted. This criterion then guides the assessment designer to consider the environment or situation within which the

assessment is to occur as a factor in the final assessment design. Whilst it is likely that the students undertaking this training would work within an office environment on completion of their training, it was felt that with some limited re-design of the layout of the classroom environment within which they were working, as well as their ability to self-pace and work outside of the classroom, it would be possible to increase the degree of fidelity with their future work environment. This response supported the feedback of students that suggested that the facilities provided in the training classroom had not previously been either adequate or workplace relevant, and neither were they equipped with the tools that would likely be used in the workplace.

Figure 5.2: Apply the process of educational multimedia evaluation to a Training Technology Centre developed Computer Based Learning Package — Trainees will construct their own

6. Is *fidelity* required in the assessment environment? And in the assessment tools (actual or simulated)?

The issue of the *fidelity* of the learning environment and the tools deployed in it was a matter of concern raised by students on post course evaluation. The students specifically raised the issue with respect to the teaching of theoretical concepts with little or no practical application and, moreover, it was noted that when the application of that theory was expected in practice, the environment within which this occurred, and the tools with which it was expected to be achieved, were not entirely not consistent with their subsequent workplace experience. Their response was that the assessment for the previous iteration of the module made little or no attempt to situate the activity in a workplace relevant context and further it did not require the application of actual workplace tools particularly software.

This meant that it would be unclear the extent to which competent performance in the learning environment could be directly equated to competent performance in the workplace, particularly where there were competing demands which meant a need to prioritise activity, and, in addition, a different set of workplace tools might be were available.

It was determined that the redesigned assessment activity would seek to be situated within a high fidelity working environment and that the tools to be applied in practice would be the same as those applied in the workplace. So the redesigned assessment activity undertaken in relation to the module, did present a much higher degree of fidelity in terms of the actual tools used to undertake the assessment than had previously been the case.

In particular, emphasis was placed upon ensuring that the software and hardware made available to the students to complete the assessment activity were consistent with those available to them in the workplace. These principally related to the items of stationery used, as well as the software packages upon which the students designed and developed their assessment outcome. In short, the tools that the students were using within the classroom were the same as those that they would use within the work environment. Whilst it is acknowledged that the layout of computers for example, on desks around the edge of the room, was not what would be expected within an office environment, students did contrive to re-arrange other aspects of the room, principally its furniture, to suit their own preferred working arrangements.

7. Does the assessment activity require *discussion* and *feedback*?

The ability to discuss, give and receive feedback is critical to much workplace performance. The value of feedback both from a teacher and peers, as well as straightforward mentoring or guidance, from a senior colleague, is an important means by which performance may be enhanced. The assessment for the earlier version of Module 10 had offered students little or no opportunity for either discussion or feedback

Student feedback suggested that they viewed a requirement for discussion and feedback as being an integral component to successful assessment performance, thus the redesign of the assessment of Module 10 would need to establish requirement for students to offer versions of the evaluation tool that they had created to classmates who would formally review, evaluate and critique them, and provide them with feedback.

The intention of this part of the redesigned assessment activity being to demonstrate to the students the value of using colleagues as a resource in improving a product or a

performance. At the same time, and throughout the conduct of the training, the teacher/researcher was available to provide feedback.

8. Does the assessment activity require that students *collaborate*?

A strong feature of the original design for this learning module had been the limited opportunity provided to the students to collaborate with one another. In particular, the design of the module and the assessment activities had encouraged students to work individually on their own assessment outcomes in a more competitive assessment environment that further detracted from the value of seeking collaboration from colleagues, and provided for little or no opportunity for collaboration.

As the opportunity for student's collaboration was considered to be an integral component of an authentic assessment performance, within this revised module, as highlighted above in Figure 5.3, the intention, from the outset, was that the re-design would overtly seek to break away from this more traditional process of having students develop assessment outcomes in isolation, comparing them with one another on completion of the assessment activity, and instead, seek to build-in the opportunity, even requirement, for collaboration during the development of assessment outcomes, as an integral and necessary part of that development process.

Figure 5.3: Process of multimedia evaluation model assessment activity

Description of how the critical questions were applied in the design and structure of the learning outcomes and assessment criteria of Module 10

What follows is a description as to how each of the individual critical elements was applied in the re-design of the module's learning outcomes and assessment criteria. A

description is given as to how each of the learning outcomes was interpreted, followed by a consideration of the design decisions made during the re-design of each of the assessment criteria that relate to them.

Learning Outcome 1.0 — Explain educational multimedia evaluation models

The purpose of this Learning Outcome was to enable students to gain an insight and general understanding in the field of educational multimedia, prior to their consideration of the elements that they would be seeking to evaluate. Firstly, students were provided with a description into the factors that determine a training product as being educational multimedia, as well as information on some of the major approaches adopted in the delivery of educational content by means of electronic, computer-based, and multimedia. Finally, they were provided with information as to the main types of evidence that should be gathered for the evaluation of an educational multimedia product.

Assessment Criteria 1.1 — Define educational multimedia

As described above, and as Benazet (2001, p.23) notes, ‘a resource is a multimedia one when it calls simultaneously upon different sensory registers and when it generates an interactivity between the learner and the artefact, made up in the majority of the cases of a data-processing device’. In this view, interactivity is fundamental and distinguishes the multimedia from the audio-visual.

Within this section the students were briefly asked to consider the ways in which advances in technology have led to changes in the design and delivery methodologies available in the field of education. In particular, they were asked to note the changes that had occurred from the early to mid 1980s, and of those, their attention was drawn to

the advent of the personal computer with a graphical user interface (GUI) capability. At this stage within the module, they also gave consideration as to how trainers and educators had begun to seek more complex and sophisticated ways of using this technology to deliver and manage education and training.

In addition to the historical perspective students were presented with a short consideration as to the underlying complexity that is often inherent in educational multimedia. This notion is supported by Avellis and Finkelstein (2002, p. 121), for whom, 'educational multimedia has an intrinsic complexity', in that it can be viewed as both, 'software running on a computer and an educational resource.' For the design of this module students began to consider these dimensions so that an educational multimedia package could begin to be viewed from its pedagogical perspective as a learning resource and also in terms of its technical implementation. In this way it was hoped that the students would begin to acquire an understanding of the complexity of the courseware under evaluation.

The students were provided with information on the factors that influence the design and development of educational multimedia; in particular, by considering the range of factors and elements to be utilised. Thus in order to be able to evaluate an educational multimedia package, students needed to know what those factors and elements were. Students were thus given information on informed principal factors and elements including those of; instructional design, the choice of media selected by the designers, the educational requirements incumbent upon the package, its interface design and the structure of the learning content underpinning the module. Students were also informed that they would need to give consideration as to how each of the these factors and elements fitted within the package in its entirety, that is, what design and development

compromises or adaptations that may have been made to accommodate each element. Students were asked to consider that, as Pham (1998, p.107) states, 'good educational multimedia must not lose sight of the educational objective while taking advantage of what advances in technology can offer.'

Finally, within this section, the student's attention was drawn to the crucial role played by evaluation, or critical analysis, in the ongoing enhancement of new and existing packages. It was noted that in the same way that newly-trained classroom teachers and instructors must learn to review the quality of face to face teaching provided by them and their colleagues, so Computer Based Learning Practitioners, or Army instructional designers, must be able to critically reflect upon the educational multimedia content which they design, use and review.

With reference to the application of the critical elements within assessment criterion 1.1, perhaps the most prevalent was that of *challenge* in that this assessment criteria presented to each of the students the requirement to ascertain a body of knowledge that they would have to draw upon both in the summative assessment for this module, but also in the subsequent practice in the workplace. Similarly, it was necessary, for the purpose of this first assessment criterion, that the students be capable of the *transfer* of the knowledge that they were obtaining from a more general educational perspective into its specific application in the critical review of educational multimedia content. Finally, this assessment criterion required that the students both *discussed* this knowledge with one another as a means of testing their understanding and embedding the knowledge more deeply, and by means of *feedback* were able to describe their understanding to both one another and the teacher.

Assessment Criteria 1.2 — Outline the main approaches in the delivery of educational multimedia

Within this section students were encouraged to give brief consideration to the some of the broad theoretical perspectives that were likely to have influenced the designers and developers of educational multimedia. In short, students were expected to be able to consider the importance of the theoretical educational approach that underpinned the educational design decisions made by the designer, prior to evaluating an educational resource.

Whilst there have been many attempts to classify the ways in which people learn, students again noted three main current approaches to structure in educational design, namely, Instructivist or traditional instructional design, constructivist and free access. It was demonstrated that from these broad classifications, a range of theories had been evolved, each of which provided a methodology for the design of learning content.

The three approached were amongst those considered to be the most prevalent within the educational design of the multimedia packages likely to be evaluated by these students. As well as being those most usually represented, they were also the ones most easily recognised by students.

Learning outcome

Within the area entitled *Learning Outcome*, the work of Gagné was presented as, one of the foremost theorists in the field of instructional design. Students noted that it was Gagné who presented the educational perspective that it was possible to both identify and measure learning by means of establishing learning outcomes. Students were informed that Gagné believed that it was possible to identify these learning outcomes in a body of knowledge or skill, and that these learning outcomes could then be

systematically identified and measured in terms of knowledge by performing tasks and sub-tasks, organised in a hierarchical fashion. This was an important theoretical perspective for these particular students to understand, as much of the training with which they would have been familiar, both in the earlier stages of their military careers in the training they undertook, and latterly as those responsible for developing and maintaining such training, would have been premised upon the work of Gagné, which had had a particularly profound impact upon the development of post World War II military training doctrine.

Cognitive

Under the heading of *Cognitive,* students were next presented with the cognitive approach and its particular focus upon knowledge acquisition as being a cumulative process, as opposed to a necessarily hierarchical one. Students were expected to note the impact that this perspective could have upon changing the emphasis of instructional design away from that of providing instruction in a teaching mode of delivery, with the focus being increasingly placed on the student in terms of the role of learning. Moreover, the students were expected to be able to identify the impact that such an educational theory might have upon the design of educational multimedia content, thus by understanding that, if knowledge acquisition becomes a learner-paced organic and cumulative process, as opposed to a hierarchical one, then the instructional designer of educational multimedia might be better able to employ the tools available within the multimedia environment, for example hyper-linking. In this instance, students reflected upon the value of hyper-linking noting that it could become a means of enabling students to evolve their own cumulative pathways through the education experience, as opposed to being forced to follow a pre-set, hierarchical menu of content. For students schooled on the learning outcome orientated approach of Gagné, this would perhaps

have been considered to have been a more radical design consideration to account for in evaluating an educational multimedia package.

Affective

Finally, under the heading of *Affective* students considered the role of considering the affective approach to instructional design. They noted that affective approaches place a greater emphasis on establishing the emotional and psychological aspects of the learner's responses. Thus the measurement of these aspects of a learner's response begins to establish the learner's motivation to learn, which will likely, impact upon the overall success of the learning experience. The attention of the students was drawn to the motivational theory of Keller and his ARCS Model as an example of this approach in action. This model, based on learner's attention, the relevance of the content to the learner and the level of confidence with which the learner approaches the content, as well as identifying the degree of a learner's satisfaction acquired from completion of the learning. For the purposes of the military training with which these students were familiar, it was important that they be able to both understand and consider the role that motivation would play in effective educational design, prior to being properly able to evaluate it.

In achieving the expectations set by assessment criteria 1.2 the students were again required to demonstrate the ability to *transfer* the learning that that they achieved across the range of knowledge areas as they brought them together to form their own understanding of the main approaches that underpin the design of educational multimedia. In this respect, it is considered also that the achievement of this assessment criterion did require that *metacognition* was demonstrated, most specifically by means of a critical reflection upon the different approaches to the design of educational

multimedia that they were offered. Finally, as with the first assessment criteria, assessment criteria 1.2 also required that the students engaged in *discussion* both with one another as well as the teacher to be able to further reflect upon their understanding.

Assessment Criteria 1.3 — Describe the types of evidence to be gathered from the evaluation of educational multimedia

In the final element of this first section of the module, students considered the range of questions that needed to be evolved, before they would be able to conduct a rigorous evaluation of a piece of educational multimedia. They noted Alexander and Hedberg's assertion (1994, p. 235) that, 'the four main approaches for evaluation were, objective-based, decision-based, value-based and naturalistic'. In addition, they were given material that provided them with information to assist them in determining the questions that they would need to consider in evaluating educational multimedia content. They noted that, firstly they would have to consider the types of evidence available to the evaluator, and that in order to do this they must be aware of the individual features of an educational multimedia package that was to be evaluated, as noted by Squires (1997), 'predictive evaluation is concerned with the assessment of the quality and potential uses of the software application, prior to its use by students'.

Students were informed that the main areas to be evaluated included both the, 'objectives and content of the package,' (Pham, 1998, p. 107) the quality of the interactivity, the package's overall attraction, and the teaching strategies that it employed as well as it technical reliability. Furthermore, before being able to evaluate these areas, students were encouraged to give particular consideration to the areas of interface design and navigation functionality.

In the consideration of interface, they were informed of the importance of ensuring that they reviewed the ease with which students can apply the interface to achieve its intended outcome, the degree of user satisfaction that a student could attain in the use of the courseware, the relative degree of ease with which the student could learn to use and operate the courseware and finally, that they consider whether the interface itself encouraged and enhanced more effective and efficient performance of the intended outcomes.

With respect to the field of navigation students were encouraged to give consideration to the ease with which trainees could access the relevant materials to either obtain knowledge or perform a task, the degree to which the learning content assisted students to acquire a deeper understanding as to how various concepts may interrelate, the ways in which the navigation employed may have been used to enhance student interaction with the materials and, finally, whether the navigation methodology enhanced creativity or encouraged a higher level of student participation and engagement with the course material.

In undertaking this assessment criterion, the students were *challenged* in the expectation to understand a large array of integrated elements, all of which combine in the creation of an engaging learning experience. The educational multimedia experience is by definition a complex one, with a large number of interdependencies that revolve around not just knowledge management and design, but also factors such as use of colour, sound, and decisions as to the best medias to use to support a given educational requirement. In many respects this criterion embodied much of the learning outcomes that the students had achieved during the previous nine modules of the Computer Based Learning Practitioners Course which has sought to build, in stages, their understanding

of the use of multimedia in educational delivery, and so enable them to bridge the practical requirement from being predominantly face to face classroom practitioners to becoming educational multimedia instructional designers.

This criterion was also to be fundamental in the student's final summative assessment requirement, namely, the development and application of their own educational multimedia evaluation *product*. In addition, as with the two previous assessment criteria, students were again required to be able to *transfer* this knowledge from the domain of the classroom and the theoretical to its use in the applied context of the evaluation of a multimedia learning package. Finally, through the acquisition of this assessment criterion the students were expected to both demonstrate *metacognition* means of critical reflection upon the various types of evidence that they would review in evaluating such a package and the requirement to embed this learning by means of both *discussion* and *feedback*.

Learning Outcome 2.0 — Outline the structure of an educational multimedia report

The second learning outcome was provide as a means to enable students to begin the process of considering the ways in which evaluation data, once collected, could be collated and structured into a report. Students were encouraged to consider that any report as well as reporting purely on the design factor noted, should also seek to set out means by which any data collected might be used to enhance the quality of a particular piece of educational multimedia software.

Students were also informed that despite existence of the four orientations of Alexander and Hedberg (1994), for their purposes, the evaluation of educational multimedia might

be considered to be within two distinct categories, serving two distinct purposes, namely formative and summative evaluation.

Assessment Criteria 2.1 — List the approaches and methods for evaluating elements of an educational multimedia report

Within this section, students began to become familiar with a range of the approaches that could be applied in the evaluation of educational multimedia. These were largely based upon what Albion (1999, p.1) referred to as, 'Useability inspection', as the generic name that he applied to the method of evaluating based upon the, 'considered judgements', of evaluators. Also described was Pham's assertion (1998, p.107) that, 'one major concern [was] to evaluate how well different aspects of the product serve to achieve the objectives of knowledge acquisition.'

Students were provided with a brief outline of the four major historical orientations in the field of evaluation of educational multimedia, as set out by Alexander and Hedberg (1994, p. 19). This was a means to enable them to understand that the role for which they were being trained had evolved over a period of time, thus in consideration of the task of performing an evaluation of an educational multimedia package there was a broader historical context within which their thinking might fit.

The four major historical orientations in the field of educational multimedia evaluation described to the students were, firstly, the 1940s Objective-based or summative evaluation, where the evaluation was predominantly objective-based. Students were informed that the intention of this method of evaluation was to determine how successfully the educational objective had been met through the application of multimedia in the training situation.

Secondly, students considered the 1970s, decision-based or formative evaluation, and were informed that this was the period within which educational multimedia evaluation became decision-based. Here the students noted that it was at this time that the more modern purpose of evaluation began to appear with the search to establish a processes for the evaluation of multimedia at all stages during the development process. Thus, students noted the way in which this adoption of a decision-based approach led to a situation whereby evaluation could occur during the design and development process and not subsequent to it.

Next, students where asked to consider the 1980s naturalistic method of evaluation as a formative approach to evaluation. This approach was outlined for students by means of a description as to how from the 1980s evaluation of educational multimedia had become increasingly more naturalistic in its approach. Students noted the ways in which, in this regard, the focus of evaluation had shifted in line with the changing educational theory of the time, to the achievement of the goals to be achieved in the education process, and a subsequent consideration as to whether, in educational terms, the goals had been worth achieving.

Finally, students considered the 1990s 'holistic/integrated' evaluation approach for use in the delivery of both formative and summative evaluations. Students were informed that from the middle of the 1990s, the evaluation of educational multimedia had shifted its focus to become both more holistic and more integrated. During this era, recognition occurred that evaluation could be a process of value at both the formative design stage as well as the summative or testing stage of development. From mid 1990s onwards evaluators began to become more systematic in the evaluation methodology that they applied, usually working with clearly defined questions and goals.

This assessment criterion *challenged* the students to make a shift in their thinking from the consideration of educational multimedia from an instructional design perspective, to now considering the ways in which they might construct a tool that would enable the evaluation of that learning experience. Again, the students were expected to be able to *transfer* the learning that had occurred throughout this course to this means by demonstrating the ability to apply their understanding of the theoretical perspective on evaluation, in practice, a process which itself would require of them a degree of *metacognition* via their critical reflection upon the knowledge that they had acquired.

Assessment Criteria 2.2 — Identify the elements to be evaluated

The students next, under the heading of 'Identify the elements to be evaluated', were expected to give consideration of the actual elements that they would evaluate within an educational multimedia package.

So students were informed that, prior to commencing any review of an educational multimedia package, they, as the evaluator, must seek to establish the overall educational purpose or intent of the package, and then consider the audience at which it has been aimed. Once these factors had been considered then they were informed that they would be ready to commence the macro evaluation of the package. As Pham (1998, p.107) put it, 'one major concern is to evaluate how well different aspects of the product serve to achieve the objectives of knowledge acquisition.'

This process of considering the evaluation from first a macro and then a micro perspective is consistent with the work of Avellis and Finkelstein (2002, p.2), who assert that the evaluator should, 'group the characteristics of multimedia educational software under four evaluation categories: educational features of the software;

technical features; aspects relating to the ease of use (useability); and, aspects relating to the content.' At the same time, students were informed of the importance of context or environment, as Reeves (1992, p. 48) notes, 'learning is highly tuned to the situation in which it takes place.'

Assessment Criteria 2.3 — Describe the components of a revision plan

Next, prior to undertaking the final summative assessment activity, students are requested to give consideration as to what the individual component parts that would comprise a revision plan might consist of. This element of the teaching phase is mediated by means of a class discussion.

In considering this area of the evaluation process students are reminded as Pyne (1994, p.34) states that, 'we evaluate educational multimedia to improve the program during its development stage (formative), to facilitate comparison with competing programs and to contribute to the general knowledge about effective design.'

In the completion of this assessment criterion the students are *challenged,* by means of the use of *discussion* and *feedback* that is used as the primary means of conveying this component of the content, with seeking to be able to take the ideas and knowledge that they have gained from the previous assessment criteria up to this point and to think about how they could use that knowledge, in a workplace context, to inform the revision of a multimedia package under evaluation. This activity in itself requires that the students demonstrate *metacognition* by means of their ability to critically reflect, self-assess and evaluate.

Learning Outcome 3.0 — Apply the process of educational multimedia evaluation to a Training Technology Centre developed computer based learning product

Finally, the students undertake the final learning outcome, Learning Outcome 3.0. This component provides the summative assessment activity, by which the students will be assessed. As the means by which the outcomes previous two learning outcomes was assessed, this component required that students be able to identify those elements relevant to the determination of successful educational multimedia when considering the value of a particular package, and subsequently, be able to report them in an appropriate manner. In terms of authenticity, in particular, with reference to the critical elements under consideration, it was important that the outcome of this activity would be a product that could be suited for use within a work environment. This was largely due to the fact that on successful completion of this the final module in their course, students would be returning to their respective work environments with the expectation that they be able to produce such outcomes. It was equally important for the students that the report that they produced would be provided in a manner that would not only enable the evaluation evidence gathered to be used to enhance the particular product under review, but that it should also used to enhance the overall design and develop processes of subsequent computer based learning packages.

Thus students were required to construct and then apply their own educational multimedia evaluation tool to an Army designed and developed educational multimedia package. On completion, they were, after consultation with at least two other course members, expected to revise their evaluation tool, and then re-apply it to another Army educational multimedia package.

Assessment Criteria 3.1 — Summative Assessment — Identify the aims of the computer based learning package

The summative assessment activity was comprised of four parts.

Firstly, students were, as stated above, to construct their own educational multimedia evaluation tool, and then apply it to an Army educational multimedia package. Secondly, they had to critically review and evaluate the performance of their educational multimedia evaluation tool in the light of feedback from at least two other course members, and produce a (100–250 word) critique.

The most obvious implementation of the critical elements in practice, for the purposes of this assessment criterion was the requirement for the student to provide a *product* as the final assessment outcome. This activity in itself also required that they be able to *transfer* the learning that had occurred by means of the demonstration of a skill, and again, this would not have been successfully achieved had they not demonstrated *metacognition* by means of critical reflection and self-assessment or evaluation. The outcome of this first two parts of this assessment activity were also assessed by a *client* in the form of the teacher, and they were required to demonstrate a degree of *accuracy* in the product that they had developed in that it had to be able to perform the task for which it had been built.

There was also a requirement that a degree of *fidelity* be displayed in the assessment environment itself, as described in more detail below, as the tools used in the development of the assessment outcome were those that would be used in the workplace. This assessment activity also required that students engaged in *discussion* with one another and sought *feedback*. The opportunity for the students to *collaborate* in undertaking this activity was also available.

Assessment Criteria 3.2 — Report of the value of the computer based learning package

On completion, and as the third part of this summative activity, students had to use the peer review and evaluation data acquired to revise their evaluation tool, which they would then re-apply, in its revised format, to another Army educational multimedia package. Fourthly, and finally, the students concluded with a critical report (250–500 words) of the package that they had reviewed, based upon the outcomes of the instrument that they had developed, noting both positive and negative aspects and features, and making recommendations as to how it could be improved.

As with assessment criteria 3.1, criteria 3.2 continued with the requirement for the production of a *product*, although for the purposes of this criterion the product was as developed previously but refined by means of *discussion* and *feedback.* The assessment activity also continued the requirement for the students to demonstrate *metacognition*, again by means of critical reflection, self-assessment and this time, evaluation of the first iteration of the evaluation tool that they had developed, to ensure that it would continue to represent as an reliable or *accurate* indicator of the quality of the multimedia package that it was to evaluate.

The role of formative assessment in the redesign of the module

Throughout the two days that the students undertook this module they were formatively assessed both formally and informally by a range of means. With respect to the formal requirement for a formative assessment, students were regularly quizzed and took part in teacher-led class discussions on the knowledge content that they had just acquired. As progress through this package was self-paced, the timing of these discussions was crucial, especially if the formative evaluation by the teacher was used as an indicator of

understanding. In this respect, it was important to ensure that the students who were progressing more slowly had reached significant points in the module before expecting them to take part in a discussion relevant to that content that would enable them to demonstrate the knowledge that they had acquired. It is worth noting here, that as well as the information that the teacher obtained, this teacher-led discussion process, according to student feedback, also enabled the students to test and confirm their own levels of knowledge acquisition and understanding, and so make their own formative judgements as to their relative degree of progress.

The increased requirement for formative assessment occurred in the early stages of the summative assessment activity. In this respect, as the students began to embark on the summative assessment they continued to seek feedback from both the teacher and one another. This guided their early progress in the activity. A similar process is likely to occur in a workplace where, once given a task to complete, an employee might seek further feedback from a supervisor to crystallise their understanding of the requirements. However, in the context of the classroom, it is likely that students felt somewhat more secure in expressing what might, in the workplace context, be construed as a lack of knowledge. Thus, on the basis of the responses received from students it appears that they felt better able to take a risk and, perhaps, make a mistake. In this respect, the completion of the summative assessment of this module, with its reliance upon discussion and collaboration, presented the students with an opportunity to establish a cultural practice in the classroom that, if they carried it forward into the workplace, would further improve their professional performance.

The application of the elements to the learning environment

Finally in this chapter consideration is given to the design of the physical learning environment within which this module was conducted and how well it related to the critical elements, and how they were applied in the learning environment itself.

As described in Chapter 3, the learning environment was comprised of a classroom, with individual personal computers arranged around three of the four walls within the room. In addition, each of the students was provided with an individual desk set in the middle of the room facing towards the front of the classroom, in a more traditional classroom design and each student also brought with them an individual laptop computer with which they had been issued at the commencement of the course.

The room was set up initially with the desks presented individually in three rows of two desks per row, and each student selected where they wanted to sit. Whilst the room was pre-set up for the initial introductory parts of the module's delivery, students were able to subsequently move furniture if they desired to better accommodate their work practices.

It was intended from the outset that the method of implementation of this module of training would as far as it was possible, seek to provide a high degree of *fidelity* in terms of both the environment and the tools that the students used, to reflect the likely work environment within which they would subsequently find themselves employed. To this end, and as described in more detail in Chapter 3, they represented a mixture of ranks from Army Lieutenants to Captains and a Royal Australian Navy Lieutenant (Army Captain equivalent). Whilst this did accurately represent the likely range of peers with whom they would need to *collaborate* within the workplace.

The notion of *fidelity* of environment was continued where students were provided with daily timings that reflected the typical defence, office-based, work day. However, what did represent a unique experience for the students was the fact that, unlike their normal military work environment, they were not required to juggle multiple tasks and deadlines, and instead, in the training environment, they had the comparatively unique experience, from a work perspective, of only having to fulfil one expectation at a time. This meant that although they were expected to communicate and collaborate with one another in the completion of this activity, they were, by and large, free from the distractions of competing priorities, and the factors that go along with this such as telephone calls and personal interruptions from staff often working on the other priorities.

As can be seen, an attempt was made to ensure that a high degree of *fidelity* was provided within the student's work environment, and, in addition, the nature of the *challenge* that confronted the students, that is the requirement to produce a product as the final assessment outcome was also consistent with a workplace expectation. In particular, the design of the module was such that it required that they be able to demonstrate the ability to successfully *transfer* knowledge obtained during previous modules of the course, as well as this module and further to this, that they demonstrate the *metacognitive* ability to both critically reflect upon the outcomes produced and, in addition, be able to *discuss* and *collaborate* with one another as required.

Conclusion

As can be seen the redesign of Module 10 sought to ensure that the feedback received from students on evaluation of earlier courses was allied to the critical elements in the

development of a learning module, and more specifically its assessment activities, that could be considered as being more authentic for the purposes of the students undertaking it.

It is also worth re-iterating at this point that it was by means of the application of these critical questions, as evolved in the previous chapters, that this module, and more particularly its assessment, was revised to ensure that it provided a more authentic learning experience for the students. In this respect, the first of the subordinate research questions: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic task?* was addressed, in that a series of specific characteristics, or critical questions for the purposes of this study, had been identified from the literature, reviewed and revised by a variety of methods, including expert review, and then applied to the design of this module of learning as the specific characteristics that would facilitate the design and assessment of this particular complex authentic task.

The next chapter, Chapter 6 focuses upon the implementation of the module within the learning environment and the subsequent analysis of the data obtained and seeks to answer the second of the subordinate research questions: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment?*

This is undertaken by means of a description of the students' experience of the module's delivery via the collation, analysis and interpretation of notes made during its delivery, as well as the observation notes made by the researcher during course delivery, and video tape that was taken at the same time. In addition, the analysis of the student's responses is reviewed.

CHAPTER 6: LEARNERS' RESPONSES TO AUTHENTIC ASSESSMENT

In Chapter 5, the learning environment that instantiated a theory-based solution to the problem of determining the extent to which authentic assessment may provide an effective model for task design and assessment was described in detail. The re-design of this multimedia learning environment, based on the critical elements was described, together with the development process used to create the module. In addition, the first of the subordinate research questions: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic task?* was addressed, in that a series of specific characteristics, or critical questions, were applied to the design of the module as characteristics that would facilitate the design and assessment of this authentic task.

The purpose of this next chapter, Chapter 6, is to describe the implementation of Module 10 and the analysis of all data collected to answer the research questions. Thus Chapter 6 focuses firstly upon the implementation of the module within the learning environment, and the subsequent analysis of the data obtained, and seeks to answer the second of the subordinate research questions: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment?*

The chapter begins with a description of the student's experience of the module from data obtained via the collation, analysis and interpretation of notes made during its delivery, as well as the observation notes made by the researcher during course delivery, and video tape that was taken at the same time. In addition, the analysis of the students' responses is reviewed.

Learning Module Implementation

As described in Chapter 3, Module 10 of the Australian Army's *Computer Based Learning Practitioners course* was delivered as an educational multimedia learning experience within a supervised classroom environment. Module 10 was the final module delivered, over a period of two days, at the end of a two week residential course. The module was designed to prepare students for their future role as a Computer Based Learning Practitioner in the Australian Defence Force, responsible for the design and development of defence educational multimedia training packages. As described in Chapter 5 each of those trained would commence a posting in such a position in the year following completion of the course. The Australian Defence Force's Computer Based Learning Practitioners Course which is the focus of this research was delivered towards the end of the calendar year prior to the January commencement of the next posting cycle.

Learning Module Evaluation and Analysis

Method of implementation

The following sections provide an analysis of the data which comprises:

1. Data gathered from six face-to-face interviews conducted with the students who undertook Module 10 of this course, 'Evaluating Educational Multimedia'. The students have been identified by the pseudonyms John, Mike, Brad, Billy, Kathy and Brenda. The interviews were conducted at the completion of this module.

2. Observation notes made by the researcher at key points as the students undertook the module. These key points were identified as being those where events occurred which were considered to be significant to the impact that the re-designed module was having on the student's ability to learn.

3. Video recordings of student activity throughout the duration of the module's conduct which lent support to the researcher's observation notes.

4. The collated responses to the evaluation questionnaires that the students completed at the conclusion of the module.

The method of analysis

The constant comparative method was the principal method of analysis used for the conduct of this research. Glaser and Strauss (1967, p.101) outline two general approaches to the analysis of qualitative data. In the first approach an analyst codes data first and then analyses it in order to test a hypothesis. In the second, a researcher wishes only to generate theoretical ideas, and thus, must continually re-design and review theories against the unfolding data set. Glaser and Strauss (1967, p. 102) also offer a third option, an analytic procedure of *constant comparison*, which combines the explicit coding of the first approach and the on-going theory development of the second.

Glaser (1965, p. 439), describes the constant comparative method as occurring in four stages, namely, '(1) comparing incidents applicable to each category, (2) integrating categories and their properties, (3) delimiting the theory, and (4) writing the theory'. In the first of these stages, that of comparing incidents applicable to each category, the analyst begins by coding each incident in the data in as many categories of analysis as

possible. The second stage, integrating categories and their properties, is a process which begins in a small way, by means of the use of short memos or conferences. However, it is at this stage that the constant comparison changes from comparison of incident to incident to a comparison of incident with properties of the category which resulted from the initial comparison of incidents. In the third stage, delimiting the theory, delimiting features of the constant comparative method are set in to prevent the task of comparing data from becoming overwhelming. This delimiting occurs at both the theory level as well with the original list of categories proposed for coding. By the fourth and final stage, writing theory the researcher will have coded data, a series of memos and a theory. Thus, in Glaser's words (1965, p. 443), 'the discussions in the memos provide the content behind the categories, which are the major themes of the theory'.

In summary, the constant comparative method was selected as it was considered to be the best mechanism by which the researcher would be able to analyse the range of data selected and determine the most meaningful outcomes.

Applying the constant comparative method

For the purposes of this research a comparison of the incidents applicable to each category commenced with the review of the data sources: noted observations of student performance, responses given within face to face interviews and notes made on review of video taped elements of the assessment delivery as well as the questionnaires and subsequent interviews undertaken at the end of the course. Next, the aggregation of the observations from these various data sources enabled the researcher to commence the process of comparison change from 'comparison of incident to incident to a comparison

of incident with properties of the category which resulted from the initial comparison of incidents' (Glaser, 1965, p.440).

The completion of this stage of data aggregation and categorisation, led to the third of Glaser's stages, that of delimiting the theory. Here, the theory was solidified. Thus 'major modifications became fewer' (Glaser, 1965, p. 441) and a set of properties began to emerge from the independent unrelated incidents which had been coded. Finally, came the writing of the theory, as by this stage a coded data set existed along with a series of memos and a theory.

Analysis of responses

As outlined in Chapter 1, the problem under examination has been that of determining the extent to which authentic assessment may provide an effective model for task design and assessment. In order to establish whether it is possible to provide a solution to this problem it became necessary to determine whether it were possible to harness the principles of authentic activity to guide the design, development and application of more meaningful, more authentic, assessment activities and thus establish the extent to which authenticity may provide an effective model for task design and assessment.

In order to answer this question, the following two subordinate questions were addressed:

1. What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?

2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?

The first of these questions has been addressed in detail in Chapter 5; this chapter now deals with analysis of the data obtained and considers how it relates to the second these research questions:

2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?

The intention here is to consider how each of the elements performed from the student perspective, when applied within the context of this module.

Researcher's observation on students' responses by data source

The following section outlines the responses received from the student interviews and other data sources.

Interview

At the completion of the training, and when the completed written questionnaires had been received and collated, trainees undertook a face-to-face, one-on-one interview, conducted by the researcher using the responses that they had provided to the student feedback questionnaire (Appendix 2) and described in more detail in Chapter 5, and using the interview questionnaire (Appendix 3).

During this process a number of common themes began to arise in the form of common response given by a number of students. These responses included such considerations

as a view that, 'authenticity was predicated upon the relative degree of accuracy that an assessment could re-produce in comparison to an actual workplace environment', in support of this particular example, one student even noted that 'an authentic assessment is one which accurately reflects that which had to occur in the workplace'. In general, however, the observations made concerning student's responses towards authenticity indicated that they saw it as being more predisposed towards workplace, or vocational training as opposed to the more academic environment of the school or university.

There was also some discussion on the role of assessment as the means to assess the information/knowledge/skills that had just been acquired, without concern as to how it would relate to subsequent workplace performance. For at least one of the students, authenticity in assessment related to the requirement of the assessment tool to 'evaluate the specific skills and knowledge that had been taught'.

Observation

The observations made during the conduct of the module's delivery were again consistent with the responses that were provided via the written questionnaire and at interview. At the commencement of the activity some of the students had initially been reluctant to participate, but that they had quickly become involved in the more active components of the module and had been eager to share their views with fellow students. Conversely, others had fully and enthusiastically participated in the learning module from the outset and established themselves as both reflective and thoughtful individual practitioners in action. A good example of this was the student who although generally positive, tended to make most comments towards the end of discussions once the issue had been considered, and this students' approach was to reflect upon the views

expressed by fellow students. Another student also had previous experience in the development of evaluation tools and their application in a learning context, and this was apparent from the speed with which this student had been able to achieve an initial first draft evaluation tool.

However, it was also observed that, whilst this previous experience had proved helpful to the student in the initial tool development, it did, to some degree, make it harder for them to accept the formative critique of colleagues.

The review of the observation notes made during the classroom activity indicated a high degree of consistency between noted comments at interview and observed classroom performance.

Video

The video observations again demonstrated comparatively high levels of actual involvement within the more interactive components of this lesson, particularly where it involved the opportunity to exchange views, ideas and opinions with fellow students. Though many of the students from both their written response and at interview expressed some degree of concern, verging upon criticism, of the principles of applying authenticity to a learning event such as this, the review of their actions during the course of the learning often proved to be largely inconsistent with the expressed views.

As described in Chapter 3, video was recorded during particular activities undertaken by the students. In all a total of 3 hours of video footage was recorded for subsequent analysis, and to support noted observations made by the researcher at the time.

Video observation provided further support to the importance that some of the students had placed upon full engagement during the classroom phase of this module. One was positive in the development of their own module and quick to seek and provide constructive feedback to others. In many respects, this student treated the classroom environment in the way that it had been intended to be used, that is as an office or workplace environment, and viewed the expected outcome as a work task, as opposed to just a measure of his competence. It was clear from the performance of such students, and the feedback that they provided, that they had viewed themselves as working to produce an authentic assessment outcome, the evaluation tool, which they would seek to use subsequently in the work place, and at the same time gain exposure to working in such a collaborative environment.

Observations made on consideration of the video evidence of part of the classroom activities also supported the comments made that whilst they had been fully engaged in all activities; the students had sought to be critically reflective of the process in which they were involved. Some of the feedback suggested that students were both immersed in the activities within the classroom, whilst at the same time seeking to reflect upon their value to themselves as learners.

Notes on student's performance made on observation during the delivery of the module

Before analysing the observations made by the researcher during the module's delivery, it should be re-affirmed that the names attributed to the students in the following sections are pseudonyms, used to ensure that actual student's cannot have specific remarks made attributed to them and that student confidentiality is maintained.

The observations made during the conduct of the modules delivery were consistent with the responses that were provided via the written questionnaire and at interview. During the course of the activity, John had been initially reluctant to participate, but had quickly become involved in the more active components of the lesson and had been eager to share his views with fellow students. Conversely, Mike had fully and enthusiastically participated in the conduct of this learning module and was a reflective and thoughtful individual in action, and though positive, many of the comments that he made tended to be towards the end of discussions once he had considered and reflected upon the views expressed by fellow students. Mike also had previous experience in the development of evaluation tools and their application in a learning context, and this was noted in the speed with which he was able to achieve an initial first draft evaluation tool. However, it was also observed that, whilst this previous experience had proved positive in initial tool development, it did, to some degree, make it harder for Mike to accept the formative critique of colleagues. This seemed largely due to the fact that the draft tool that he developed had been based on accepted military practise in the design of such a tool. When faced with feedback from fellow students that challenged the considerations implicit in the design, Mike had found it more difficult to deconstruct and re-consider his design in the light of such feedback. It is considered that the behaviours represented by John and Mike were largely reflective of how they would have performed within a similar work environment. Furthermore, as the two most senior members of the group by military rank (John being a Royal Australian Navy Sub Lieutenant and Mike an Australian Army Captain); it is also possible that some of their initial reticence was born from their desire to ensure that they were not undermined by the performance of a more junior colleague.

The observation notes made upon the classroom performance of Brad indicate a high degree of consistency between his noted comments at interview and his observed classroom performance. Brad had been very quick to acknowledge the requirement of this module for a produced outcome, and sought to focus his path through the content of the module to that outcome. In particular, Brad was very conscious of the need to gain constructive feedback that would enable him to develop the evaluation tool that he had produced into a viable measuring instrument. Again, as with John and Mike, the way in which Brad was challenged to perform through this activity, and the way in which he responded to that challenge, was very likely an accurate representation as to how he would have performed in a work environment.

Again, as with most of the other students, the observations made of Billy's classroom performance were largely consistent with his stated views. Billy was happiest when working alone on his own individual outcome, and was to some extent reluctant to seek the feedback of classmates on the outcomes of his work. Billy preferred to work in a more self-reflective manner, and sought to enhance the quality of the tool that he had produced by means of personal comparison with exemplar tools that he had accessed on-line. So also Kathy's observed performance was consistent with the comments that made during interview. Kathy had sought to be very involved in the activities within the classroom, and was a willing participant in the process of considering the evaluation tools as developed by classmates, suggesting also that Kathy would be a willing participant in workplace-orientated team activities.

Brenda was also fully engaged and involved during the conduct of the classroom activity. As well as being quick to provide opinion and involve herself in discussion, she was also very keen to reflect upon the views put forward by other and, by the

comments made at interview and the notes made during observation, appears to have understood the longer-term workplace relevance to her of the degree of authenticity designed into this module.

Notes on student's performance made on researcher review of the video content recorded during the delivery of the module

The video observations of John again demonstrated his comparatively high level of actual involvement within the more interactive components of this lesson, particularly where it involved the opportunity to exchange views, ideas and opinions with fellow students. John had both within his performance and, subsequently, at interview expressed some degree of concern, verging upon criticism of the principles of applying authenticity to a learning event such as this, however, and his actions during the course of the learning proved to be largely inconsistent with his expressed views.

The analysis of the video content of Mike provided little additional insight, other than that already noted above, as to the views of Mike or how far his methodology of working through the content of this module further supported, or otherwise, the value of the critical elements central to the module's design.

Video observation did lend further support to the importance that Brad placed upon full engagement during the classroom phase of this module. Brad was positive in the development of his own module and quick to seek and provide constructive feedback to others. In many respects, Brad treated the classroom environment in the way that was intended, that is as an office or workplace environment, and viewed the expected outcome as a work task, as opposed to just a measure of his competence. It was clear from his performance and the feedback that he provided that he viewed himself as

working to produce an authentic assessment outcome, the evaluation tool, which he would seek to use subsequently in the work place, and at the same time gain exposure to working in such a collaborative environment.

Billy was particularly involved during class discussions on all related topics and was very forthright in expressing his views and opinions on the values of such evaluation tools and the ways in which they could best be applied. Billy was also very constructive in the nature and tone with which he provided formative feedback to classmates. It also appeared when observed, that the quality of the feedback provided was generally of a good level.

Observations made on consideration of the video evidence of part of the classroom activities further supports the comments made by Kathy, in that whilst she was fully engaged in all activities, she did seek to be critically reflective of the process in which he was involved. The feedback that Kathy provided suggested that she was both immersed in the activities within the classroom, whilst at the same time seeking to reflect upon their value to herself as a learner.

Brenda was a whole-hearted participant in the broader discussion that occurred during the classroom phase of this experience. She was also able, upon subsequent consideration, to reflect upon the content of the discussions that occurred and at the same time think about those outcomes in a more holistic way in terms of both the overall impact as well as the value of authenticity in educational assessment.

The student's response to the critical questions

The following section considers the responses received from the students and the feedback that they provided on a critical element by element (or critical question by question) perspective, and seeks to relate this feedback to the second of the subordinate research questions: *How do students' respond to tasks designed to incorporate the characteristics of authentic assessment?*

1. To what extent does the assessment activity *challenge* the assessed student?

Whilst *challenge* had been included as a critical element of authentic assessment on the basis that to be authentic, an assessment activity must be challenging to the student undertaking it, it was acknowledged that, if the activity was deemed by the student to be either too easy, or too obvious, for example the simple repetition of recently acquired knowledge, then it was less likely to be consistent with the sometimes less than clear outcome required within the workplace.

With reference to the relative degree of *challenge* implicit within this module, the respondents were generally in agreement that they had felt a degree of *challenge* in undertaking the activity. However, the level or degree of the challenge felt varied between the students.

For some, notably, John, they had felt that that overall the actual *challenge* presented had been low, although the value of the knowledge obtained from it had been high. Additionally, some had felt that the expected assessment outcome, or challenge, had, been too close to the overall training provided throughout the course, and came after the students had undergone a previous two weeks of training in the field of educational

multimedia, during which time they had had the opportunity to peruse a range of evaluation tools. John stated that he had been able to complete the final assessment outcome by means of cutting and pasting from evaluation tools that he had previously seen during earlier parts of the course, as opposed to creating a new one of his own. However, he did also note that he considered that the activity had been high value on the basis that it provided for a better appreciation of the evaluation instruments. Whilst John's feedback with regard to the area of challenge was received and noted with interest, it was also discussed with John, that the production of a new evaluation instrument, by means of utilising what were considered to be the better aspects of previous tools, was not in itself considered to be negative, provided that there was sufficient difference in the overall instrument developed to demonstrate the understanding applied by the student in its design.

Mike supported this feedback and stated that the *challenge* presented to trainees had been what he described as, '7 out of 10'. However he did acknowledge that the *challenge* inherent to this activity had largely been determined by the fact that this was the first time that he had designed and deployed an evaluation tool of his own for this purpose. Furthermore, he noted that whilst he had previously conducted evaluations of lesson content, he had never had to create and then apply his own evaluation tool to do so.

Whilst overall the response seemed to suggest the *challenge* had not been considered to be high, some, most notably Billy, did also note that the assessment had required 'thought and research' to achieve an outcome, and, therefore, as it was not 'looking for a straightforward answer' it had been 'quite challenging'.

But not unexpectedly, *challenge* experienced was discovered to be an entirely relative concept, largely dependent upon the particular prior skills, knowledge, and to some extent degree of confidence, that the individual student brings to the activity concerned. This is a notion supported in this instance by Brad who stated that the process employed in designing and then constructing an evaluation instrument, alongside the subsequent review of another evaluation tool constructed by a peer, meant that, in his opinion this module and its assessment had presented him with, 'a high degree of challenge'. He further noted that this process had required the application of the knowledge that he had acquired in both the creation of the evaluation tool and the critical review of one built by a peer, thus it had, 'enabled a full synthesis of knowledge and cognitive skills'.

However, consistent with *challenge* is the idea that to be more authentic as an assessment, students should be required to demonstrate the ability to synthesise from the range of skills and knowledge that they have acquired those necessary for the completion of a specific outcome or outcomes. This is a process requiring both analysis of the task, and the subsequent selection of the appropriate response, as in real world situations and tasks the necessary response will often require the synthesis of a range of skills and information into the formulation of a potentially correct response.

This synthesis was supported by a number of the students, most notably, Brad, Kathy and Brenda who believed that the *challenge* presented to students had been good, and, as stated by Kathy, had, 'seemed to integrate a large amount of the previously received course material, as taught within other modules of the course'. For Kathy, as with Brad, the final assessment was being viewed as the 'culmination of the whole course of study that had preceded this final module'. A view also supported by Brenda who considered that there had been a sufficient *challenge* presented to the students. She cited the

requirement to develop an evaluation tool and then use it to evaluate a piece of educational multimedia courseware as the principal example of the degree to which she had felt that she had been challenged.

The opinion as to the actual *challenge* presented within this activity was varied. Whilst the majority felt that they had perhaps not been sufficiently challenged by the final assessment activity, it did appear that whether they realised it or not, each of the trainees had been challenged, and each had used a range of authentic techniques to achieve the final outcome including that of reflecting upon already produced tools designed for the purpose of evaluating educational multimedia.

2. Is a *performance*, or *product*, required as a final assessment outcome?

The next critical element of authentic assessment considered is that of *performance,* or *product,* as a final assessment outcome. With respect to its application within the module, the intent had been to recognise that within the authentic workplace environment, successful performance is often, if not usually, measured by means of the production of a specific work related *performance* or *product*. Determination, or assessment, of success in the world beyond the educational environment is then often defined by the quality of the final *performance* or *product* that is developed on request. When questioned about the value of this module in terms of performance, or product as a final assessment outcome it was generally agreed by the students that the importance of producing a crafted outcome had been central to the successful completion of the module.

John and Mike had both directly commented upon the fact that a product had been expected as the final assessment outcome, and Mike noted that, as finalisation of the

module had required the completion of a product, consistency with this critical element of authentic assessment had resulted in the development of a product that would be directly applicable in his forthcoming workplace. Or as Archbald and Newman (1988, p. 12) describe state, 'students demonstrate skills and knowledge by engaging in complex performance, creating a significant product or accomplishing a complex task using higher order thinking, problem-solving and often creativity'.

This was a view generally supported by most of the students who observed that the most positive aspect of this assessment activity had been its requirement for the completion of a product that would be directly applicable. However, Billy did note that it would have been valuable to have been able to review the evaluation tools that had been developed by each of the other students at the completion of the module. Kathy also commented, as previously noted, that the ability to complete the final stage of the assessment activity, namely using the constructed evaluation tool to evaluate a piece of educational courseware, required the development of an evaluation tool, then for the purposes of the assessment of this module a product had to be viewed as an integral part of the final assessment outcome.

In conclusion, it can be seen that the students identified both the existence of the *performance or product* as the final assessment outcome and the value of having either measurement of performance or development of a product as a valid determinant of assessment outcome. What was interesting, however, and not foreseen, was the way in which some of the students viewed the development of the product as the intended outcome, as opposed to its application. In short, and in terms of application within the context of this module, it is considered that both can be viewed as having relevance in the determination of final assessment outcome.

3. Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?

The requirement to demonstrate the *transfer* of learning from other related areas of study was the next critical element of authentic assessment that was included within the design of this particular module. The ability to apply knowledge, skills or attitudes from one domain to another is often dependent upon the understanding, and application, of knowledge from other domains. In this respect, the authenticity of a learning outcome, and the assessment used to measure it, should recognise the requirement to be able to *transfer* learning that has been acquired in other domains, or elsewhere within the same domain, what Tanner (1997, p. 8), describes as, 'consistency between the assessment and the real-world application for which the learner is being prepared'.

Opinion amongst the students as to whether *transfer* of learning had been adequately or appropriately applied within the context of this module remained divided. In particular, John felt that, with the exception of the application of MS Word Skills, little *transfer* of learning had either occurred or been required, though, he did note the value of the 'personal self-reflection' that the assessment had encouraged, as an indicator of the explicit inclusion of learning *transfer* and as means of improving his overall performance outcome.

Conversely, others considered that *transfer* of learning had been seen to have occurred noting particularly that they had been required to apply learning that they had undertaken in this and in previous modules on the two week course, in order to complete the assessment outcome. Mike also noted that since the completion of the course he had applied the tool that he created as for the final assessment outcome to

evaluate and improve a piece of educational multimedia content that he was developing privately.

Some, most notably Brenda, also pointed out that a large degree of *transfer* had been required between the other modules of the Computer Based Learning Practitioners course where multimedia evaluation had been touched on, and the required assessment outcome for this particular module of that course. The general opinion was that as the requirement had been to produce an evaluation tool this must be seen to represent a realistic assessment of competence, in this respect, and as Mike stated, 'a sound performance of the assessment activity could be directly *transfer*red to workplace performance'. He did note, however, that a delay in opportunity to practise this skill within the workplace could lead to a degradation of the skill and a diminishing of the degree of any eventual *transfer* that might occur from the learning to the working environment.

Whilst at the outset the notion of *transfer* of learning was intended primarily as a means of ensuring that consideration be given to confirming that knowledge, skills and attitudes taught or required during training were explicitly measured in the design of assessment tools, some of the trainees saw another value implicit in the notion of *transfer*, namely that of *transfer* from the training environment to the workplace, consistent with Berlak's (1992, p. 25) who described transfer in terms of, 'the degree to which the assessment is related to what the learner is being prepared to do beyond the particular assessment setting'.

4. Does the assessment activity require that *metacognition*, is demonstrated, by means of critical reflection, self-assessment or evaluation?

In determining authenticity within an assessment activity, the next critical element that was designed into this module was the importance of ensuring that students would be able to apply *metacognition* by means of critically reflecting upon and self-assessing or self-evaluating the assessment outcomes that they were producing. When questioned about the module in terms of critical reflection and self-assessment or evaluation the six respondents were in full agreement that it had been utilised within this learning activity.

The students considered that both critical reflection and evaluation had been integral to the assessment activity for this Module. In particular, they noted that the use of more than a single peer reviewer had forced the students to apply *metacognition* via critical reflection in a positive manner, and that not only was it possible to obtain feedback from a fellow student but it was also valuable experience to undertake the role of critically reviewing the work of others. In addition to this, two students in particular, Mike and Brad, noted the two stage process of reflecting upon their own work initially and then reflecting on it again based on the comments received from peers, consistent with Khattri, Reeve and Kane (1998 p 14) who note that the process of *metacognition* can be developed if, 'students [are] required to explain the process by which they arrive at answers'. Some of the students, including Mike, commented that the hand over of their work to a colleague for peer review had encouraged them to more deeply and more critically self-reflect on and evaluate the work that he was handing over, to ensure both that it was of a sufficiently high standard, and that they understood the reasoning behind the design of their evaluation tool.

Another important observation was that of Kathy who noted that the drafting phase of the development of the evaluation tool had assisted in enhancing the role played by *metacognition*. She noted that it was by means of critical reflection, self-assessment and evaluation that she had not just improved the design of her own evaluation tool, but had also improved the quality of the feedback that she was able to provide to fellow students in the design of their evaluation tools, in support of this, Brenda observed that the requirement to evaluate another student's work and provide them with feedback was the best example of this in action.

It was apparent that the students had applied *metacognition* in support of the development of their own evaluation tools, as well those of fellow students. Moreover, it appears from the responses received at interview that they were overtly aware of the value of this process in ensuring a positive assessment outcome.

5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?

This element acknowledges the importance, particularly within the work environment, of ensuring that a required product or performance is accurate, or, to the required standard. In addition to this, is the value of recognising the role that environment plays in determining the ability of an individual to perform at or to a required standard.

With regard to this element the six students were generally fairly dismissive of the way that in which they felt that this element had been applied within the assessment for this particular module. Only two, Billy and Brenda, considered that the assessment activity had called for a high degree of *accuracy* in performance and that similarly high degree of fidelity of assessment environment had been either required or provided. Both also

expressed the view that they did not believe that *accuracy* of performance had been required, at least, not in terms of a workplace standard. However, Billy did provide some comment that he felt that the evaluation tool that he, in particular, had constructed had been accurate, and that, furthermore, there had, therefore, been some requirement to demonstrate accuracy.

Most of the students acknowledged that the actual determination of the level or degree of *accuracy* required was a function of that expressed in the learning outcome. In support of this Billy stated that, 'as far as the requirement for *accuracy* was concerned, particularly in the design and development of this evaluation tool, the determination of its degree of accuracy wrested on the intended learning outcome'. If the learning outcome had been that of designing an evaluation tool, he considered that the assessment had displayed a high degree of accuracy. If, on the other hand, the intended learning outcome had been to apply an evaluation tool, then he felt that the degree of accuracy displayed in this instance was low. This feedback is of interest to the assessment designer, as particularly in the area of performance, it notes that accuracy will be both context and outcome dependent, as described by Moorcroft et al (2000), where they outline the duality of the purpose designed into authentic assessment tools as not only assessment tools but also exercises through which students may be enabled or encouraged to explore their deeper understanding of a topic often as they are applying the knowledge, (p. 20).

In considering this element in detail, some, namely Mike felt that the assessment outcome had been too removed from the workplace. In his opinion, there was an insufficient requirement for the tool to be applied against an actual package or product where the feedback could be applied in the enhancement of that real world package.

This degree of simulation, in Mike's opinion, detracted from the degree of accuracy, and also impacted negatively upon the *fidelity* of the assessment environment.

6. Is *fidelity* required in the assessment environment? And the assessment tools (actual or simulated)?

The *fidelity* of the assessment tools used is the means by which a consideration is given to the tools that are provided within the assessment environment to construct the assessment product or performance. This has relevance within both the actual, as well as the virtual, assessment environment where it is vital to ensure, as far as is possible, that tools provided have a high degree of *fidelity* with those that will subsequently be found within the work environment.

The students were generally in agreement that the tools that they had applied to the completion of their own assessment activity had represented a high degree of *fidelity*. Most, such as John, noted that they considered that this assessment tool had exhibited an extremely high degree of *fidelity* and that the tools used had required an assessment outcome that he considered to be identical to that applied in the workplace, this is consistent with Berlak (1992, p. 24) for whom, 'the hallmark of authentic assessment practices is their harmony with real world circumstances'

However, he did also note that even though the level of *fidelity* of the tools used to build his assessment outcome was high, the time that had been allocated for the completion the tool had not been adequate and thus reduced the degree of authenticity in his view. However, other students noted that the ability to use tools such as the Internet, especially Google, in the classroom had ensured a high degree of fidelity as such tools would be readily available within the workplace to undertake this task.

Overall, it appears that the students considered that they were provided with a sufficient *fidelity* in the tools that they were able to use within the assessment to complete the outcome, and that from the designer's perspective, they did have ready access to the same tools that they would expect within their next work environment.

7. Does the assessment activity require *discussion* and *feedback*?

Next considered was the requirement to ensure the inclusion of *discussion* and *feedback* as a critical element in the design of an authentic assessment activity. In this regard, and consistent with actual workplace performance, it is rarely that an individual undertakes the completion of a work activity without the benefit of discussion with colleagues and the ability to receive and benefit from their feedback. As Northcote and Kendle (2000, p.8) have noted, 'it is extremely helpful to build in opportunities for feedback in assessment'

With reference to this module, the students agreed that they had observed and utilised a requirement to *discuss* and both give and receive *feedback* in undertaking the assessment activity. In particular, most agreed with the response of Mike that there had been a high level of requirement for *discussion and feedback*, both student to student, as well as, student to facilitator. Billy also stated that he considered that a fair amount of discussion had occurred, in particular, 'discussion between students', as they shared ideas on such matters as how best to deal with the assessment requirements. He also noted that the best feedback had arisen from the ability to trial the evaluation tool against a product and then analyse its performance.

Most considered that the *discussion* and *feedback* that had occurred was the central mechanism for the learning that was taking place, and noting the discussions in which

they had personally been involved, considered that they had aided the assessment process. Mike in particular noted that, he considered that the discussions that he was able to hear as a result of the configuration of the comparatively small classroom had also further enhanced the ability of the student group to benefit further from *discussion* and *feedback* and that this had played a part in their successful completion of this assessment.

The peer review component of the assessment activity was also noted as having assisted in the enhancement of a greater degree of focus in *discussion* and also an increased degree of *feedback*. However, in the opinion of Billy, this feedback process led to too much reliance upon the role of the instructor as expert, as it had been too easy to illicit instructor response and feedback during the formative stages of evaluation tool development. However, Brenda had commented that the level of feedback required, both received and given had been good, particularly where it related to the final assessment outcome. She also noted that there had been opportunity for discussion to occur throughout the conduct of the training.

Some feedback was received, however, that this critical element of the assessment's design could have been improved upon. In particular, it was felt by some that once the second draft of the evaluation package had been developed, with the benefit of reflection and peer review, it would have been of value to have included a mechanism to enable students to submit the improved evaluation tool to peers for further review. It was considered that this second phase of peer review would have enabled the review of peer progress against feedback already received and encouraged further *discussion* on *feedback* already received.

8. Does the assessment activity require that students *collaborate*?

The final critical element designed into the assessment activity of this module was that of *collaboration*. Unlike *discussion* and *feedback*, which are intended to focus more fully upon the use of feedback to improve a student's individual assessment outcome, the notion of collaboration is more a shared one, where two or more students are enabled to work collaboratively in the completion of a shared assessment outcome.

As Bruffee (1984, p. 647) asserts, 'in business and industry, and in professions…where to work is to learn or fail — collaboration is the norm'. In this context the role of the teacher becomes that of a guide while students collaborate to, 'make connections between new ideas…and prior knowledge, use language as a tool for learning, and develop language and thinking competencies' (Bayer, 1990, p. 7).

For the purposes of this module the assessment outcome was predicated upon the completion of an assessable product from each student, though it was noted that no-one sought the opportunity to *collaborate* with a fellow student in the development of a shared outcome. If then, the importance of collaboration is that it recognises within the workplace that there is very often a requirement to perform as a member of a team and that the final outcome may only be achieved through the active collaboration of a designated group, then in this regard the assessment activity might be deemed as not having performed well.

The responses of the students, when questioned about the module in terms of the ability to collaborate, were mixed. Whilst some, such as John, considered that the level of cooperation required between students and the facilitator had been of a low level, Mike conversely expressed a view that he believed that the design of the assessment activity

had required that *collaboration* occur, and drew particular attention to the phase of review, once students had received feedback from fellow students. At this point, he noted that, 'students needed to further refine the feedback received, by means of collaboration, to understand what worked and why', and conversely what did not work and, why not.

It was also generally felt that the requirement for peer review had ensured that collaboration had been able to occur during this process. However, as Billy noted, 'the requirement to collaborate' was very much, 'a function of an individual's own learning style', and that as his preferred style was a more individual one, he had sought less opportunity to collaborate. He did observe however that whilst all trainees were offered the opportunity to work collaboratively on a shared outcome, each chose to work on an individual outcome.

Some students, as represented by the feedback of Kathy, felt that with regard to the role of collaboration, the drafting process had been an effective means of encouraging collaboration; however, Kathy stated that she felt that it could have been further, 'improved upon'.

In conclusion, whilst the students recognised the overall value of collaboration to an activity such as this, the ways in which they sought to *collaborat*e with one another tended to vary on an individual basis, and as stated above, it is of interest that despite being able to collaborate on the development of a shared assessment outcome, each of the trainees ended by taking the option of producing their own individual product.

Summary of the student's response to the application of the critical questions in the redesign of Module 10

In summary, it appeared that most of the students acknowledged an authentic assessment as one which was representative or reflective of a performance that would have to occur within a workplace, which led them to the view that authentic assessment was more pre-disposed towards training or education in a vocational training context as opposed to the often more academic environment of the school or university. This is consistent with Lund, where he states (1997) for whom the value inherent in authentic assessment tasks is that they establish connections between real world experiences and school-based ideas (p. 25).

With regard to the *challenge* that had been implicit within this module, the students appeared to be of the opinion that it ranged from a medium to a high level, and whilst they had undertaken a previous two weeks of training in the field of educational multimedia, they still felt sufficiently challenged by the expected assessment outcomes required from this module, and stated that, in their opinion, it had a high value in that it had provided them with a better appreciation of educational multimedia evaluation instruments and their design.

Additionally, whilst some students claimed to have had previous experience in the conduct of evaluations, they noted that, in the main, they had not had the experience of creating and then applying an evaluation tool of their own making, thus further increasing, in their opinion the overall relevance, and thereby authenticity, of this activity.

Some disagreement had occurred between students as to whether, in their opinion, the final assessed outcome should have been considered to be a *performance* in that it required the evaluation of educational multimedia, or a *product* in that to undertake this evaluation they had had to construct a product in the form of the evaluation tool. It is worth noting however that most of the students did ultimately note the duality of the final assessment requirement as both process and outcome; performance and product.

Further debate also ensured as to whether a *transfer* of learning had occurred successfully within this assessment activity. Whilst the majority appeared to be satisfied that *transfer* had occurred, noting that necessary content had been provided to enable completion of the assessment outcome, not all students were convinced that they had been sufficiently required to ensure that the knowledge acquired in the earlier part of the module had been used in the construction of the evaluation tool.

Most students expressed the view that the requirement to produce an evaluation tool, in this instance, represented a realistic assessment of competence, within the boundaries of the desired performance. However, some concern did exist as to the lag between knowledge acquisition and knowledge application.

Areas that were considered to be critical to the determination of authenticity within assessment were those of critical reflection and evaluation. Students had noted, in particular, the use of peer review, which they felt had encouraged critical reflection to occur. Thus it had not only been possible to obtain feedback from a colleague on their own performance, but also to review the work of others. Students had also commented upon the fact that the handing over of work to a colleague for peer review had encouraged them to reflect even more critically upon the work that they were handing

over, in the main to ensure that it would be considered to be of a standard that they considered to be sufficiently high for authentic workplace application. At least one of the group had stated that this process, 'not only encouraged students to reflect more deeply upon the quality and content of their own work, but also to apply the same level of reflection on the comments received from peers'.

Another theme that emerged from the analysis of the conduct of this module was the high level of concern amongst the students to produce what they and their classmates would consider to be workable or useable outcomes.

Again, at least one student expressed concern about the requirement for accuracy, stating, as previously noted, that if the intent of the learning outcome was to, 'design a multimedia evaluation tool', then the assessment had required a high degree of accuracy, however, if the intended learning outcome had been to 'learn about evaluation', then the degree of accuracy displayed had been low.

In terms of *fidelity* it appeared that most students had considered that this assessment activity had exhibited a high degree of fidelity, as the requirement of the assessment outcome was consistent with the eventual workplace performance expectation.

Further to this, the role of *discussion* and *feedback* was considered to have been the central mechanism by which students felt that their learning had occurred. All students were broadly in agreement that the significant requirement for *discussion* and *feedback*, both student to student, as well as, student to facilitator, had been critical. The requirement for peer review was also viewed as a mechanism that had ensured a greater degree of focus in *discussion* and an increased degree of *feedback*.

However, at least one student had felt that use of *discussion* and *feedback* could have been further enhanced with a subsequent phase of reflection and peer review on completion of the second draft. It was felt that this would have enabled the subsequent review of peer progress against feedback already received, and encouraged further discussion on that feedback.

It was evident that the necessity for *collaboration* was considered to have been high. In this respect, particular attention was drawn to the phase of review undertaken as students received feedback from fellow students. Some stated that they had needed even more opportunities to *collaborate* and further refine the feedback received, than they had been provided with, so as to better understand what worked and why and conversely what did not work, and why not.

The next section will seek to build on the students' response to the individual critical elements and consider how it relates to their actual response to the assessment activity.

The students' response to the assessment activity

The overall perception was that this assessment could have made greater use of *collaboration* and, in particular, conducting the assessment in pairs or groups of three would have enhanced this experience. The assessment also seems to have been viewed by the students as a highly individualised assessment task, and one that they also felt could have been improved with the availability of sample evaluation tools as a form of template to the way in which Army currently meets this requirement. However, the students did acknowledge that overall the module and its assessment had increased their awareness of how to evaluate an educational multimedia package.

Moreover, some of the students had expressed concern that for some of the more senior students (in terms of military rank), particularly within a hierarchical work environment such as the military, the receipt of feedback from those who were, 'subordinate in rank' might be an issue, though he stressed that he had not witnessed this occurring during the delivery of this module.

Others had expressed the view that, if the assessment activity had required the actual collection of data and information by means of the application of the evaluation tools, then this could have increased the inherent degree of authenticity contained within the activity. It was also noted that the opportunity for students to have been able to consider data and draw conclusions from it as well as to make decisions and recommendations based on it could again have increased the level of authenticity of the assessment.

Further feedback received had suggested the potential for the addition of a 'post assessment' forum for further discussion after the evaluation tools had been assessed. It was also felt that such a forum could also have been used as a means for the group to review some exemplar tools.

Some feedback had been received to suggest that this assessment activity could have been made more authentic if the evaluation tool developed had been able to be applied to an actual piece of educational multimedia courseware currently under development.

However, one student, in particular, noted that the assessment activity had presented a concise and relevant assessment activity that had followed the following workflow:

Learn Theory→Create Tool→Apply Tool→Reflect on Outcome

This work flow in itself provided a neat and concise representation of the process that had been followed in implementing the design of this learning module.

Discussion

It appears that overall this assessment could have made greater use of collaboration and, in particular, conducting the assessment in pairs or groups of three could have enhanced this experience. It also appears that the assessment was viewed by the students as a highly individualised assessment task, and that they felt that the activity could have been improved if the facilitator had made available some sample evaluation tools that provided some form of template to the way in which Army currently meets this requirement. However, it was acknowledged that overall the module and its assessment had increased the group's awareness of how to evaluate an educational multimedia package.

Further to this some stated that they felt that the timeframe within which the activity had to be undertaken had impacted upon the quality of the assessment outcome, particularly in the application of the completed evaluation tool to a package, and noted that the relative inexperience of the students in the field of educational multimedia design and development was a factor that had impacted negatively on the assessment activity overall. In future, particularly if delivered by distance means the pressure that the students felt to complete the assessment outcome in the timeframe might be reduced by providing them with additional time. It should be noted however, that delivery by distance would bring additional issues for students over and above those identified in this process.

It was also noted that for some of the more senior students (in terms of military rank), particularly within a hierarchical work environment such as the military, the receipt of feedback from those that were, 'subordinate in rank' might be an issue. This is a particularly useful comment as it adds a level of sophistication to the underlying factors to be considered when applying a collaborative learning construct. Even though hierarchy is overt in the military, it had not been considered as an issue during task design. Thus within organisations or groups where such hierarchy is more covert, it is important that its existence is both acknowledged and accounted for when designing for a collaborative learning outcome.

Further comment noted that a requirement for the collection of workplace data and information by means of the application of the evaluation tools could have increased the inherent degree of authenticity contained within the activity. The opportunity that this would have presented for students to have been able to consider real workplace relevant data and draw conclusions from it as well as to make decisions and recommendations for the workplace based on that data would again have increased the level of authenticity of the assessment. Again this is a valid consideration, but the degree to which it would be possible to achieve will vary greatly based upon the context within which the learning is delivered.

Comment was also raised with respect to the potential to include a post assessment forum for further discussion within the assessment. It had been felt that on completion of the assessment outcome (the multimedia evaluation tools) it would have been valuable for students to have had a forum to provide a means for the students to review some exemplar tools. This was an interesting idea that could be implemented in future,

in both a face to face or distance mode, with minor timetabling or curriculum reorganisation.

It appeared that most of the students acknowledged an authentic assessment as one which was representative or reflective of a performance that would have to occur within a workplace, which led them to the view that authentic assessment was more pre-disposed towards training or education in a vocational training context as opposed to the often more academic environment of the school or university. In support of this it had been noted that, 'if training was the goal, then assessment needed to be authentic', however, the issue was to consider whether the same requirement existed for authenticity when the goal was education or knowledge application from a more academic or theoretical perspective. In this regard, the students involved in this activity appeared to believe very strongly that to be considered as authentic, an assessment should relate to either a workplace or performance outcome.

With regard to the degree of challenge that had been implicit within this module, the students appeared to be of the opinion that it ranged from a medium to a high level, and whilst the students had undertaken a previous four weeks of training in the field of educational multimedia, they still felt sufficiently challenged by the expected assessment outcomes required from this module, stating that, in their opinion, the expected assessment outcome from this module had a high value in that it had provided them with a better appreciation of educational multimedia evaluation instruments and their design.

Additionally, whilst some students had previously had experience in the conduct of evaluations, they noted that, in the main, they had not had the experience of creating and

then applying an evaluation tool of their own making, thus further increasing, in their opinion the overall relevance, and thereby authenticity, of this activity. One student in particular, reported that the process of designing and then constructing an evaluation instrument, alongside the subsequent review of another evaluation tool constructed by a peer, had, for him, ensured both the, 'immediate application of the acquired knowledge', as well as giving him the opportunity to practice his ability to provide a critical review of the work of a peer. In his opinion, this had required the full synthesis of both knowledge and cognitive skills, and, for him, the final assessment represented the culmination of the total course that he had undertaken, concluding with Module 10.

Some disagreement had occurred between students as to whether, in their opinion, the final assessed outcome should have been considered to be a performance, in that it required the evaluation of educational multimedia, or a product, in that to undertake this evaluation they had had to construct a product in the form of the evaluation tool. It is worth noting however that most of the students did ultimately note the duality of the final assessment requirement as both process and outcome; performance and product, one student stating that to be able to undertake the final stage of the assessment activity, namely, the evaluation of educational multimedia courseware, then the requirement for a product had to be viewed as an integral part of it.

Further debate also took ensured as to whether a transfer of learning had occurred successfully within this assessment activity. Whilst the majority appeared to be satisfied that transfer had occurred, noting that necessary content had been provided to enable completion of the assessment outcome, not all students were convinced that they had been sufficiently required to ensure that the knowledge acquired in the earlier part of the module had been used in the construction of the evaluation tool.

Most students expressed the view that the requirement to produce an evaluation tool, in this instance, represented a realistic assessment of competence, within the boundaries of the desired performance. However, some concern did exist as to the lag between knowledge acquisition and knowledge application. It was felt by some that, 'sound performance of the assessment activity, could be directly transferred to workplace performance, provided that performance occurred sufficiently soon after completion of the training'.

Areas that were considered to be critical to the determination of authenticity within assessment were those of critical reflection and evaluation. Students had noted, in particular, the use of peer review, which they felt had encouraged critical reflection to occur. Thus it had not only been possible to obtain feedback from a colleague on their own performance, but also to review the work of others. Students had also commented upon the fact that the handing over of work to a colleague for peer review had encouraged them to reflect even more critically upon the work that they were handing over, in the main to ensure that it would be considered to be of a standard that they considered to be sufficiently high for authentic workplace application. At least one of the group had stated that this process, 'not only encouraged students to reflect more deeply upon the quality and content of their own work, but also to apply the same level of reflection on the comments received from peers'.

Another theme that emerged from the analysis of the conduct of this module was the high level of concern amongst the students to produce what they and their classmates would consider to be a workable or useable outcome. Though, this observation must sit against the backdrop of some of the received student feedback that expressed concern that there had been an, 'insufficient requirement for the tools to be applied to an actual

package or product', where any feedback could have been directly applied to the improvement of a piece of actual multimedia courseware.

At least one student had expressed concern about the requirement for accuracy, and it was stated that if the intent of the learning outcome was to, 'design a multimedia evaluation tool', then the assessment had required a high degree of accuracy, however, if the intended learning outcome had been to 'learn about evaluation', then the degree of accuracy displayed had been low.

In terms of fidelity it appeared that most students had considered that this assessment activity had exhibited a high degree of fidelity, as the requirement of the assessment outcome was consistent with the eventual workplace performance expectation. Further to this, the role of discussion and feedback was considered to have been the central mechanism by which students felt that their learning that occurred. All students were broadly in agreement that the significant requirement for discussion and feedback, both student to student, as well as, student to facilitator, had been critical. The requirement for peer review was also viewed as a mechanism that had ensured a greater degree of focus in discussion and an increased degree of feedback.

However, at least one student had felt that use of discussion and feedback could have been further enhanced with a subsequent phase of reflection and peer review on completion of the second draft. It was also felt that this would have enabled the subsequent review of peer progress against feedback already received, and encouraged further discussion on feedback already received.

It was evident that the necessity for collaboration was considered to have been high. In this respect, particular attention was drawn to the phase of review undertaken as

students received feedback from fellow students. Students stated that they had needed even more opportunities to collaborate and further refine the feedback received, so as to better understand what worked and why and conversely what did not work, and why not. However at least one of the students noted that in her opinion the requirement to collaborate was very much a function of an individual's learning style, and that some would always prefer a more individualised approach.

In conclusion, most students appeared to be of the opinion that the critical elements of authentic assessment had been addressed in the assessment tool designed for this module, although there was some opinion that the assessment activity itself may have been too individualised, and could have been enhanced had students been able to work in pairs or small groups to develop the outcome. It should be noted here, however, that students were, at the commencement of the module, offered the option of working in this way, but all had declined.

Further to this it was generally felt that students might have benefited from the provision of some sample or exemplar evaluation tools, noting that other feedback had related to the comparatively tight timeframe within which the assessment outcome was to be achieved; though again, it was made clear at the commencement of the module, that there was no requirement to have the assessment outcome completed by the end of the two days of instructional time. All students did complete the assessment activity prior to the end of the second day.

A number of students had also considered that the collection of actual evaluation information that should have been used to inform an existing educational multimedia

package in design or development, and that this linkage would have increased the overall degree of authenticity implicit within the assessment.

The analysis of the feedback provided by students during the evaluation process led to the requirement for further reflection upon the structure of the critical questions. Set out below (Table 6.1) is a description of the critical questions with a summary of the student feedback received when analysing the evaluation data with regard to the performance of that critical question. This table summarises the issues raised by the students and relates them to the relevant question. It is on the basis of the information summarised within this table that the decision was made as to whether or not to revise a particular element. The impact of this feedback on the individual critical questions is outlined in Table 7.1 (Chapter 7).

Table 6.1: Student Feedback on the Critical Elements

	Critical question	***Student feedback on performance***
1.	To what extent does the assessment activity *challenge* the assessed student?	Degree of challenge implicit within the activity was 'medium' to 'high'. Activity had a 'high value' on the basis that it provided for a 'better appreciation' of educational multimedia evaluation instruments. Unlikely that many students had previously created and then applied an evaluation tool of their own making. Process of designing and then constructing an evaluation instrument, alongside the subsequent review of another evaluation tool constructed by a peer, enabled both the 'application of knowledge learned' and the ability to critically review one built by a peer.
2.	Is a *performance*, or *product*, required as a final assessment outcome?	Distinction drawn as to whether assessment was a 'performance' in the form of evaluating educational multimedia, or a 'product' in the form of the tool built to undertake the evaluation.
3.	Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?	Most considered that transfer of learning had occurred. Requirement to produce an 'evaluation tool' represented a realistic assessment of competence and measure of transfer, within the desired performance.

	Critical question	*Student feedback on performance*
4.	Does the assessment activity require that *metacognition* is demonstrated, by means of critical reflection, self-assessment or evaluation?	'Critical reflection' and 'evaluation' integral to the assessment activity for this Module. Use of peer review, which had encouraged critical reflection to occur. Handing over of work to a colleague for peer review encouraged students to critically reflect on and evaluate the work that they were handing over, to ensure that it was of a sufficiently high standard.
5.	Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?	Some students noted that they believed there was an insufficient requirement for application of the tool against an actual package or product where the feedback could be applied to the improvement of that piece of courseware.
6.	Is *fidelity* in the assessment tools required (actual or simulated)?	This assessment tool had exhibited an 'extremely' high degree of fidelity, requiring an outcome that was identical to that used in the workplace. Noted that it would have been useful to have examples of exemplar products.
7.	Does the assessment activity require *discussion* and *feedback*?	'Discussion and feedback' was the 'central mechanism' for the learning that occurred. Significant requirement for discussion and feedback, both students to student, as well as, student to facilitator. Requirement for peer review ensured a greater degree of focus in discussion and also an increased degree of feedback. Requirement for subsequent phase of reflection and peer review on completion of the second draft.
8.	Does the assessment activity require students to *collaborate*?	Level of 'collaboration required' was high. Requirement to collaborate a function of individual learning style.

The next chapter, Chapter 7, provides a discussion of the issues as raised by the students during the evaluation of the performance of these elements, as well as a detailed consideration as to the impact that this feedback had upon the overall design of the critical elements.

CHAPTER 7: DISCUSSION

This chapter concludes Phase 3 of this study, where the module was implemented and evaluated in practice to ascertain the student's response to the design of an authentic assessment utilising these critical elements or questions. It also describes the impact of this feedback on the next iteration of the critical questions. The chapter then discusses the students' responses, and considers them in the light of the research questions. Up to this point, the critical elements or critical questions under consideration had been evolved from the review of the literature in Phase 1, and subsequently revised in the light of the feedback provided by professional practitioners within the field, as well as by the selected group of expert reviewers in Phases 1and 2. This chapter will further reflect upon the student response to the use of the critical questions in the design of an authentic assessment activity as undertaken in Phase 3 where the module was implemented and evaluated in the learning environment.

The learning module had been re-designed based upon the revised critical questions that had been evolved from Phases 1 and 2 of this process. The module itself had then been developed and delivered to the students. As described in previous chapters, the students had been subjected to a process of teacher/researcher observation for the duration of the two days that they had undertaken the module and its assessment. On completion of the module and the assessment activity, the students had completed an evaluation questionnaire (Appendix 2) and had been interviewed against a pre-designed questionnaire (Appendix 3). Once all students had been interviewed and their responses collected, the teacher/researcher reviewed and collated the individual responses received from the students, analysed and coded the notes taken during observation and produced a transcript of the dialogue of those elements of the classroom activity that had been

video taped. This transcript, and the additional notes taken on review of the video content, were also coded and collated alongside the outcomes of the interview scripts.

At the conclusion of this the review of the feedback, a table of the critical questions (Table 6.1, Chapter 6) was developed as a means of collating the student feedback, and further considering the critical questions in the light of the outcomes of the observations, interviews and reviews of the video material of the students as previously outlined. This process is developed further in Table 7.1, where the design of the critical questions is again revisited and reflected upon this time on the basis of the student's response.

The fourth, and final phase, is that of establishing the next iteration of the critical questions that can be provided as a revised framework for application by the designer or developer of assessment activity.

In reflecting upon the relative value of the critical elements and questions used in the re-design of this module, as determinants of authentic assessment activity, it has been important to establish the individual value of each of those elements as components in determining the extent to which overall this authentic assessment has provided an effective model for task design and assessment.

What follows is a consideration of the impact of the students' response upon each of the original research questions posed. Overall, as will be seen, whilst the critical questions themselves appeared to be appropriate determinants of authenticity in assessment, scope did yet exist for some further, albeit limited, revision.

Research questions — Data analysis

The intention of this research has been to identify and then codify the principles of authentic activity into an applicable framework that could be used to guide the design, development and application of a more meaningful, more authentic, assessment activity, in effect to establish whether *authentic assessment* could provide an effective model for task design and assessment. The answer to this question has been sought by means of establishing answers to the following two subordinate questions:

1. What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?

2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?

This section considers how the student's responded to each of the critical elements, in the light of these research questions, and sets out what changes, if any, were made to the critical questions in the light of that response.

The response of the students, as determined in the analysis of the data in Chapter 6 is set out below in Table 7.1.

Table 7.1: Consideration of the student's responses with reference to the research questions

	Critical question	*Subordinate research question*	*Student response*
1.	To what extent does the assessment activity challenge the assessed student?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic	• assessment activity must be challenging to the student undertaking • activity had been high value on the basis that it provided for a better appreciation of the evaluation instruments

	Critical question	*Subordinate research question*	*Student response*
		tasks?	• assessment had required 'thought and research' • students required to demonstrate the ability to synthesise from the range of skills and knowledge that they have acquired
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• Students felt a degree of challenge in undertaking the activity • challenge inherent to this activity largely been determined by the fact that this was the first time that they had designed and deployed an evaluation tool for this purpose • necessary response required the synthesis of a range of skills and information into the formulation of a potentially correct response
2.	Is a performance, or product, required as a final assessment outcome?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?	• performance measured by means of the production of a specific work related performance or product • success in the world beyond the educational environment is then often defined by the quality of the final performance or product that is developed on request • the existence of the product as the final assessment outcome and the value of having measurement of development of a product as a valid determinant of assessment outcome
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• the importance of producing a crafted outcome had been central to the successful completion of the module • development of the product as the intended outcome, as opposed to its application
3.	Does the assessment activity require that transfer of learning has occurred, by means of demonstration of skill?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?	• ability to apply knowledge, skills or attitudes from one domain to another is often dependent upon the understanding, and application, of knowledge from other domains • delay in opportunity to practise this skill within the workplace could lead to a degradation of the skill and a diminishing of the degree of any eventual transfer that might occur from the learning to the working environment
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• Opinion amongst the students as to whether transfer of learning had been adequately or appropriately applied within the context of this module remained divided • transfer from the training environment to the workplace would occur
4.	Does the assessment activity require that metacognition, is demonstrated, by means of critical reflection, self-assessment or evaluation?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?	• able to apply metacognition by means of critically reflecting upon and self-assessing or self-evaluating the assessment outcomes that they were producing
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• students considered that both critical reflection and evaluation had been integral to the assessment activity for this Module • noted the two stage process of reflecting upon their own work initially and then reflecting on it

	Critical question	*Subordinate research question*	*Student response*
			again based on the comments received from peers • hand over of work to a colleague for peer review had encouraged deeper and more critical self-reflection and evaluation of the work handed over, to ensure that it was of a sufficiently high standard • critical reflection, self-assessment and evaluation had improved the quality of the feedback students were able to provide to each other • students felt that it wasn't necessary to describe the different ways in which metacognition might be represented
5.	Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?	• importance, particularly within the work environment, of ensuring that a required product or performance is accurate, or, to the required standard • role that environment plays in determining the ability of an individual to perform at or to a required standard • accuracy be both context and outcome dependent
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• students dismissive of the way that in which they felt that this element had been applied within the assessment for this particular module • students considered that the actual determination of the level or degree of accuracy required was a function of that expressed in the learning outcome • the assessment outcome had been too removed from the workplace
6.	Is fidelity required in the assessment environment? And the assessment tools (actual or simulated)?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?	• Consideration was given to the fidelity of the tools that are provided within the assessment environment
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• tools that they had applied to the completion of their own assessment activity had represented a high degree of fidelity • time that had been allocated for the completion the tool had not been adequate and thus reduced the degree of authenticity
7.	Does the assessment activity require discussion and feedback?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?	• rarely that an individual undertakes the completion of a work activity without the benefit of discussion with colleagues and the ability to receive and benefit from their feedback • peer review component of the assessment activity assisted in the enhancement of a greater degree of focus in discussion and also an increased degree of feedback
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• students had observed and utilised a requirement to discuss and both give and receive feedback in undertaking the assessment activity • high level of requirement for discussion and feedback, both student to student, as well as, student to facilitator • discussion and feedback was the central

	Critical question	*Subordinate research question*	*Student response*
			mechanism for the learning that was taking place • discussions had aided the assessment process • peer review component of the assessment activity assisted in the enhancement of a greater degree of focus in discussion and also an increased degree of feedback • would have been of value to have included a mechanism to enable students to submit the improved evaluation tool to peers for further review
8.	Does the assessment activity require that students collaborate?	**1.** What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?	• collaboration is more a shared one, where two or more students are enabled to work collaboratively in the completion of a shared assessment outcome • the role of the teacher becomes that of a guide while students collaborate to, 'make connections between new ideas…and prior knowledge • importance of collaboration is that it recognises within the workplace that there is very often a requirement to perform as a member of a team and that the final outcome may only be achieved through the active collaboration of a designated group • the requirement for peer review had ensured that collaboration had been able to occur during this process • ways in which they sought to collaborate with one another tended to vary on an individual basis
		2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?	• no-students sought the opportunity to collaborate with a fellow student in the development of a shared outcome • students recognised the overall value of collaboration to an activity such as this

1. To what extent does the assessment activity *challenge* the assessed student?

In the consideration of the first question: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks*, and in the context of this first question, students seemed satisfied that this question was appropriate. Moreover, there was general agreement from the students as to the value of challenge in an assessment activity. This is particularly noteworthy as review of the observation notes made during the early stages of the module's delivery

indicated that a number of the student's had expressed a degree of anxiety as to their ability to complete performance requirement, or challenge implicit in the design of this assessment activity.

The consensus of the students that undertook this module and its assessment was that an assessment activity must be challenging to the student undertaking it and that in this regard this module had been of value on in that in challenging them, it had also provided them with a more thorough appreciation of the value of the evaluation instrument. Additionally, they concluded that the assessment had required both 'thought and research' of them, and they also noted that they had been challenged to demonstrate their ability to synthesise from the range of skills and knowledge that they had acquired

In the consideration of the second question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, the student's response, particularly on evaluation at the completion of the activity, indicated an overall positive response to the *challenge* with which they had been confronted.

Whilst noting that they had acknowledged a degree of challenge in undertaking the activity, they also recorded that they felt that the challenge inherent in this activity had largely been determined by the fact that this was the first time that they had designed and deployed an evaluation tool for this purpose, however, they also agreed that the response required by this assessment had meant the synthesis of a range of skills and information to provide the final assessment outcome.

On the basis of feedback received, it was decided that there was a requirement to make a minor amendment to this question and as a result the adjective *assessed* was deleted as from the student perspective as it appeared superfluous.

Thus in the case of the first critical question, the student response led to the following amendment to the text of the question:

Original critical question 1: To what extent does the assessment activity *challenge* the assessed student?

Revised critical question 1: To what extent does the assessment activity *challenge* the student?

2. Is a *performance* or *product* required as a final assessment outcome?

In the consideration of the first question: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks*, students expressed a degree of concern as to whether the expected outcome of the assessment activity was in fact a *performance* or a *product*. In this regard they stated that the assessment could be viewed as either the *performance* required to complete the evaluation tool, or it could be equally measured by means of just the assessment of the tool itself. However, they did agree that the existence of the product as the final assessment outcome, as well as the ability to measure performance during its development where both valid determinants of assessment outcome

This concern to differentiate between *performance* and *product* is made more apparent when considered in the context of the second question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, where the student's response, suggested that, in their opinion the consideration of *product* or *performance* was not required in a stand alone question, and it would have been better suited as a factor in determining any learning transfer that may have occurred.

Whilst the student's perspective was noted here, it was decided that this question would not, particularly in the light of the feedback set out in the next paragraph be further amended, instead, this feedback would be considered against a subsequent application of these critical questions, or delivery of Module 10.

Critical question 2: Is a *performance* or *product* required as a final assessment outcome?

3. Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?

With respect to the first question: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks*, in the context of this third critical question, the students response indicated that they considered that the ability to apply knowledge, skills or attitudes from one domain to another is often dependent upon the understanding, and application, of knowledge from other domains. They also noted that they felt that a delay in opportunity to practise a skill within the workplace could lead to a degradation of the skill and a diminishing of the degree of any eventual transfer that might occur from the learning to the working environment

In the consideration of the second research question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, the opinion of the students as to whether transfer of learning had been adequately or appropriately applied within the context of this module remained divided, though they did feel that for the purposes of this assessment that transfer from the training environment to the workplace would occur.

As with the second critical question, it was decided that the student response to this critical question did indicate a requirement for further reflection and review in a subsequent application of the questions or Module 10, but for the purposes of this study there was not sufficient data to warrant the amendment to the current wording of this question.

Critical question 3: Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?

4. Does the assessment activity require that *metacognition* is demonstrated by means of critical reflection, self-assessment or evaluation?

With reference to the first question: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks,* students appeared to be of the view that the inclusion of a requirement to demonstrate critical reflection, self-assessment and evaluation was important in the consideration of the authenticity of the module and its assessment, and that with reference to the assessment activity that they had undertaken for Module 10, that they had been able to apply metacognition by means of critically reflecting upon and self-assessing or self-evaluating the assessment outcomes that they were producing

However, concern was expressed by the students, in the consideration of the second question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, particularly about defining the ways in which *metacognition* might be undertaken.

In particular, students had considered that both critical reflection and evaluation had been integral to the assessment activity for this Module, and they acknowledged the two stage process required for reflecting initially upon their own work and then reflecting on it again based on the comments received from peers. They also felt that the process of handing over their work to a colleague for peer review had encouraged deeper and more critical self-reflection and evaluation of the work handed over, thus critical reflection, self-assessment and evaluation had improved the quality of the feedback students were able to provide to each other

In conclusion, the students where generally positive about the fourth critical question but felt that it could be revised to remove the suggestions as to the ways in which *metacognition* may be demonstrated, the concern being that listing a few, may be misinterpreted as being exclusive of other ways in which they may demonstrate metacognition. In the case of the fourth critical question, the student response led to the following amendment to the text of the question:

Original critical question 4*: Does the assessment activity require that* metacognition *is demonstrated, by means of critical reflection, self-assessment or evaluation?*

Revised critical question 4: Does the assessment activity demonstrate *metacognition*?

5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?

On review of the student feedback for this critical question, and in the light of the first research question: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks*, students expressed

support for the notion that to be authentic an assessment outcome should be applied in practice and therefore, in effect, be assessed by a *client.* Noting the importance, particularly within the work environment, of ensuring that a required product or performance is accurate or produced to the required standard.

They also picked up on the role that environment plays in determining the ability of an individual to perform at or to a required standard, particularly in consideration of the degree of its fidelity to the actual workplace.

In consideration of the second research question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, it became apparent that a number of the students had felt that their had been an insufficient requirement in the undertaking of this assessment activity to actually apply the outcome of the assessment, that is the evaluation tool, in the workplace against a current educational multimedia package in development, which would have enabled a *client* the opportunity to assess the *product* or *performance* and in this respect the assessment outcome had been too removed from the workplace.

The students felt that the tool developed should have been used within the real workplace, for the improvement of an actual piece of courseware in development. On the basis of this feedback some scope did exist for the wording of the question to seek to further strengthen the link to workplace, therefore, making it more authentic. Whilst it had not originally been intended that the outcomes of the student's assessment would be directly used in the workplace, it was noted that they might be used by students subsequently in the workplace, on completion of the course. However, it was decided that, at this stage, no further amendment to this question would be made, but that further

consideration would be given to strengthening this workplace link in subsequent applications of these critical questions or on subsequent delivery of Module 10.

Critical question 5: Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?

6. Is *fidelity* required in the assessment environment? And the assessment tools (actual or simulated)?

The requirement for fidelity in the assessment environment had arisen as a result of feedback at expert review. In considering the response to this critical question against the first of the research questions: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic task*, it appeared, particularly from observations made as to the ways in which students utilised the environment and interacted with the tools provided within it, that the *fidelity* of both the assessment environment as well as the tools, was considered by them to be an important consideration when undertaking this activity.

In considering the second research question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, students noted that there had been a high level of expectation for accuracy in the overall assessment outcome, that the assessment tool that they had developed had been expected to demonstrate *fidelity*. In fact, they noted that the tool that they had been required to build was to all intents and purposes, identical to that which they would have been expected to have applied in the workplace. The comments made by students as regards fidelity of the tools were that they would have liked to have been provided with exemplar tools from which they felt

that they could have learnt, and that in a workplace, they would likely have had access to such exemplars. However what was of interest in the responses received from students is the almost complete lack of comment regarding the fidelity of the assessment environment. In some respects this might be taken as indicative that the *fidelity* of the environment was sufficient and therefore, did not require further comment, however, this is an area that will need specific attention in the subsequent application of these questions on a subsequent delivery of Module 10, thus it was felt that at this point the student response to the sixth question did not require further amendment to the wording of the question.

Critical question 6: Is *fidelity* required in the assessment environment? And the assessment tools (actual or simulated)?

7. Does the assessment activity require *discussion* and *feedback*?

In considering the first question: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks*, it was noted that the review of the student's interview responses, as well as of the notes made at observation, revealed that students placed a high degree of importance on the requirement to ensure that *discussion* and *feedback* were included as critical questions in authentic assessment design. In particular, they noted that it is rare in a work environment that an individual undertakes the completion of a work activity without the benefit of discussion with colleagues and the ability to receive and benefit from their feedback, and in this respect the peer review component of the assessment activity assisted in the enhancement of a greater degree of focus in discussion and also an increased degree of feedback

With respect to the second question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, one student actually listed *discussion* and *feedback* as being the, 'central mechanism' by which the learning was enabled to occur. Notes taken on observation reveal that this *feedback* and *discussion* had occurred at both the 'student to student' as well as the 'student to facilitator' level.

In considering the phrasing of this question, students reported that they placed a high value on the role that peer review held within the undertaking of this activity and that *discussion* and *feedback* opportunities had provided the means to enable this to occur. This feedback supports the notion that both *discussion* and *feedback*, whether that is 'student to student', 'student to facilitator' or 'facilitator to student', are important questions in the consideration of an authentic workplace performance. In this respect, whilst no further amendment of this question was considered necessary at this stage, it was noted that further consideration may need to be given as to how best to ensure that the opportunity for student review could properly be allowed for in the design of the authentic assessment activity.

Critical question 7: Does the assessment activity require *discussion* and *feedback*?

8. Does the assessment activity require that students *collaborate*?

With respect to the first research question under consideration: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks*, each of the students indicated that they considered *collaboration* to be an integral component to successful workplace performance. Within the context of

this module they felt that it was the role of the teacher to become that of a guide while students collaborate to, ‘make connections between new ideas…and prior knowledge.

In another respect the importance of collaboration is that it recognises within the workplace that there is very often a requirement to perform as a member of a team and that the final outcome may only be achieved through the active collaboration of a designated group and the requirement for peer review had ensured that collaboration had been able to occur during this process, although they noted also that the ways in which a student may seek to collaborate will vary on an individual basis

In considering the second research question: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment*, it was noted that, whether their performance was reviewed from interview feedback or in terms of observation all appeared to place great value on the opportunities presented to *collaborate* on the completion of this activity, although it is noteworthy that none of the students sought the opportunity to collaborate with a fellow student in the development of a shared outcome.

Thus whilst collaboration was viewed as important, at least one student reported that the degree or extent of collaboration evidenced by an individual student may, to some extent be a function of an individual preferred style of learning, and that this should be reflected better in the questions under review. In this respect, the student’s individual response suggested that the design of the eighth critical question could be revised to accommodate the fact that, whilst *collaboration* may be viewed as critical, the degree to which that *collaboration* occurs should somehow take into account the individual student’s learning style.

Whilst this response was noted, it was decided that in the light of this comment only being made by a single student, there was no current requirement for any further amendment to the wording of this question.

Critical question 8: Does the assessment activity require that students *collaborate*?

Overall, the responses received from the students, both under observation and during evaluation and interview, suggest that from their perspective at least, firstly, that it is possible to establish the principles of authentic activity, and secondly, it is further possible to design them into a framework that can be applied to guide the design, development and application of a more meaningful, more authentic, assessment activity, thus, in this instance using the principles of *authentic assessment* to provide an effective model for task design and assessment.

Table 7.2 (below) provides a summary of the comments received from students on a critical question by critical question basis together with a description of any amendments made to the questions in the light of that feedback. It is noted that it is only the first and fourth of the critical questions that have been amended (shown highlighted and emboldened below) in the light of this feedback, but the comments regarding the potential future design of many of the other questions is raised as an area for further research in Chapter 8.

Table 7.2: Student Feedback on the Critical Questions

	Critical question prior to student feedback	***Student feedback on performance***		***Critical question subsequent to student feedback***
1.	To what extent does the assessment activity *challenge* the assessed student?	Degree of challenge implicit within the activity was 'medium' to 'high'. Activity had a 'high value' on the basis that it provided for a 'better appreciation' of educational multimedia evaluation instruments. Unlikely that many students had previously created and then applied an evaluation tool of their own making. Process of designing and then constructing an evaluation instrument, alongside the subsequent review of another evaluation tool constructed by a peer, enabled both the 'application of knowledge learned' and the ability to critically review one built by a peer.	**1.**	**To what extent does the assessment activity *challenge* the student?**
2.	Is a *performance*, or *product*, required as a final assessment outcome?	Distinction drawn as to whether assessment was a 'performance' in the form of evaluating educational multimedia, or a 'product' in the form of the tool built to undertake the evaluation.	2.	Is a *performance*, or *product*, required as a final assessment outcome?
3.	Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?	Most considered that transfer of learning had occurred. Requirement to produce an 'evaluation tool' represented a realistic assessment of competence and measure of transfer, within the desired performance.	3.	Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?
4.	Does the assessment activity require that *metacognition* is demonstrated, by means of critical reflection, self-assessment or evaluation?	'Critical reflection' and 'evaluation' integral to the assessment activity for this Module. Use of peer review, which had encouraged critical reflection to occur. Handing over of work to a colleague for peer review encouraged students to critically reflect on and evaluate the work that they were handing over, to ensure that it was of a sufficiently high standard.	**4.**	**Does the assessment activity demonstrate *metacognition*?**
5.	Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?	Some students noted that they believed there was an insufficient requirement for application of the tool against an actual package or product where the feedback could be applied to the improvement of that piece of courseware.	5.	Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?

	Critical question prior to student feedback	***Student feedback on performance***		***Critical question subsequent to student feedback***
6.	Is *fidelity* in the assessment tools required (actual or simulated)?	This assessment tool had exhibited an 'extremely' high degree of fidelity, requiring an outcome that was identical to that used in the workplace. Noted that it would have been useful to have examples of exemplar products.	6.	Is *fidelity* in the assessment tools required (actual or simulated)?
7.	Does the assessment activity require *discussion* and *feedback*?	'Discussion and feedback' was the 'central mechanism' for the learning that occurred. Significant requirement for discussion and feedback, both students to student, as well as, student to facilitator. Requirement for peer review ensured a greater degree of focus in discussion and also an increased degree of feedback. Requirement for subsequent phase of reflection and peer review on completion of the second draft.	7.	Does the assessment activity require *discussion* and *feedback*?
8.	Does the assessment activity require students to *collaborate*?	Level of 'collaboration required' was high. Requirement to collaborate a function of individual learning style.	8.	Does the assessment activity require students to *collaborate*?

Summary of student response and impact on the critical questions

The overall response of the students to these critical questions was positive although some amendments were made in the light of this feedback, in particular, to critical questions 1 and 4. However, in general, the responses received from the students appeared to be consistent with the feedback received from professional colleagues and expert reviewers during the development of earlier iterations of the critical questions, namely that the students both understood and acknowledged the value of this deliberate attempt to design an increasingly authentic assessment activity. With the exceptions of the two minor amendments made to critical questions 1 and 4 it also appears that they

accepted the value of the individual critical questions as determinants of authenticity in the context of this re-design.

The next and final chapter, Chapter 8, will set out the studies conclusions and consider some of the limitations inherent in this study as well as making recommendations for future research.

CHAPTER 8: CONCLUSION

Introduction

This chapter summarises the processes undertaken within the proceeding chapters, considers the conclusions reached and seeks to describe some of the limitations that should be taken into account when considering the outcomes of this research. It also reflects upon some of the ways in which this research may be further developed from this stage.

Summary and review of process

The specific problem under consideration for the purposes of this study related to the field of assessment. More specifically, as detailed in Chapter 4, it was based upon a requirement identified from the outcome of a number of post course evaluation activities where it was apparent that students transitioning from the Australian Army's Computer Based Learning Practitioners Course, into the workplace, seemed to consistently present without the necessary pre-requisite skills required to perform the role successfully. Furthermore, they would also often possess insufficient confidence in the skills that they obtained from this course.

One means of addressing this problem was to review, revise and re-design the final module — Module 10 — Evaluating Educational Multimedia, of this course, in line with a more authentic approach to the educational design of the course itself, but more specifically to ensure that the summative assessment activity was designed to provide an accurate determination of the student's suitability to commence performance of the role in the workplace.

Thus the problem under consideration was that of determining the extent to which authentic assessment might provide an effective model for task design and assessment.

In order to establish whether it would be possible to provide such an improvement in performance, it was decided to determine whether it was possible to harness the principles of authentic activity to guide the design, development and application of a more meaningful, more authentic, assessment activity. If it was, then the next step would be to establish the extent to which authentic assessment could provide an effective model for task design and assessment.

To determine the answer to this question, it was necessary to address the following questions:

1. What are the specific characteristics of authentic assessment that facilitate the design and assessment of complex and authentic tasks?

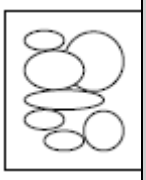
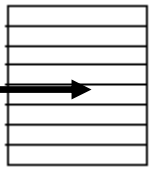
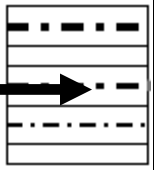
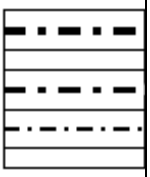
2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?

This study has sought to establish from the literature the critical elements of authentic assessment, and to develop those elements into a framework. Subsequent to this, expert analysis and feedback has been utilised to enhance the design of the elements within that framework. The framework has then been used in the design of an assessment activity within an item of courseware which has be applied to a student population in order to determine how that student body might react to that design. The responses of those students to the application of that re-designed learning module were recorded and reviewed on completion.

The study itself has been conducted in four phases. Set out below is a conceptual framework (as described in Chapter Three, Figure 3.2) of the stages of the design based research process applied in this study, together with a brief description of the activities undertaken in each of the four phases.

Table 8.1: Stages of the design-based research process in this study

Stages of the design based research process in this study

Chapter	***1**	**2**	***3**	**4**	**4**	**5**	**6**	***7**	**8**
Data		General literature to explore the problem		Discrete body of literature relevant to the defined problem	Interviews	Analysis of initial module assessment and implications for change based on the elements	Student observations Interviews Video material for analysis		
Process		Exploration of the problem Discussion with practitioners		Data mining of specific relevant studies and practitioner feedback	Further practitioner consultation and expert review	Use of expert reviewed elements to re-design module assessment	Implementation and evaluation of the module		Presentation and dissemination of key ideas
Framework of elements (the evolving product)		Key concepts – no elements 		Draft elements 	Elements reviewed by experts 	As before, now applied by researcher 	Modified elements based on evaluation 		Published dimensions 
Framework version				Version 1 (see Table 4.2)	Version 2 (see Table 4.3)		Version 3 (see Table 7.2)		
Phase in Reeves' design based research approach (see Table 3.1)		Phase 1: Analysis of practical problems by researchers and practitioners in collaboration		Phase 2: Development of solutions informed by existing design principles and technological innovations			Phase 3: Iterative cycles of testing and refinement of solutions in practice		Phase 4: Reflections to produce revised design elements and enhance solution implementation
Researcher phases aligned to design based research		Phase 1: Exploration of the problem		Phase 2: Development of a solution			Phase 3: Implementation and evaluation		Phase 4: Presentation of findings
				Phase 2.1: Development of draft elements to guide a solution to the problem	Phase 2.2: Further practitioner consultation and expert review of the draft elements	Phase 2.3: The application of the elements in the re-design of a learning module			

Chapter *1 – Introduces and frames the study, Chapter *3 describes research methodology and Chapter *7 discusses the answers to the research questions

PHASE 1: Exploration of the problem

The initial exploration of the problem had required the detailed analysis of the literature to determine the extent to which current designers had been able to gauge the degree of authenticity within an assessment activity. On completion, a series of discussions were held with a numbers of colleagues and fellow educators from both the defence and civilian sectors were held over a period of several months.

In the light of these discussions, an extensive review of the literature was undertaken to determine whether the critical elements of an authentic assessment activity, as derived from literature review, could be presented into a single set of criteria that could be applied to the design and implementation of an assessment task. The research identified that a number of educators had offered a range of factors, criteria and elements to consider depending on the context. It was decided to identify the most common of these individual factors and to synthesize them into a single cogent framework for the context of this problem. This was undertaken in Phase 2 of the research.

PHASE 2: Development of a solution

Phase 2 of the research consisted of three key activities:

1. The development of the draft elements to guide a solution to the problem

2. Further practitioner consultation and expert review of the draft elements, and

3. The application of the elements in the re-design of a learning module.

Development of the draft elements to guide a solution to the problem

Using a grounded approach a number of critical elements were established and used to construct the initial framework. A number of sources, including refereed papers, research studies and reports were identified and these papers were reviewed in detail. The principles and guidelines established within them were listed and grouped, and from these groups, eight critical elements were evolved to form a guiding framework for the study.

Further practitioner consultation and expert review

Next expert review of the critical elements was sought; firstly, and during initial development of the principles, a series of discussions was held on each with a number of practitioners in the field of education. By means of this iterative process the list of critical elements was determined and evolved. Secondly, and when this list was sufficiently developed, feedback was sought from three selected experts in the field each of whom had undertaken to act in the role of expert reviewer.

Experts considered the critical elements in terms of their value overall as a framework, as well as reflecting upon each of the elements on an individual basis. At this stage the experts began to deconstruct each of the elements in turn, and consider them on their own merits. Finally, they were asked to provide feedback that could be used to improve the elements. The framework of draft elements was revised in the light of that feedback.

The application of the elements in the re-design of a learning module

This revised framework was then used to guide the design and development of a training module within the Army's *Computer Based Learning Practitioners Course*.

The selected module, Module 10 — *Evaluate Educational Multimedia* had been delivered to a previous course the year before and it was reported on evaluation the design and delivery of this module had impacted negatively upon the student's perceived levels of confidence in actually undertaking this task within the workplace.

PHASE 3: Implementation and evaluation

During Phase 3 the re-designed and re-developed Module 10 was implemented and evaluated in practice as the final module delivered to students over a period of two days at the end of the intensive two week residential course.

The course was intended as a means of training a group of Australian Defence Force military staff in the basic principles of designing and developing training courseware for delivery by means of educational multimedia. Each of the students was to commence a posting within Army and Navy Training Command as Computer Based Learning Practitioners within the year following completion of this course.

Whilst Module 10 was designed and intended to be used subsequently as self-paced delivery at distance, on this occasion, it was delivered in a more formal classroom setting, and the revised framework of the critical elements was used as the basis for the design and development of this module and its summative assessment.

During module delivery, and on its completion, a range of data formats were collected for subsequent analysis for the purpose of this study. Whilst the training was being conducted the students were observed by the researcher and extensive notes were made as to the ways in which the students interacted with both the material as well as each other. Subsequently, these notes were collected, collated and analysed. Also during the

delivery, various aspects were videoed for subsequent analysis. On completion, students completed a written questionnaire. These written responses were collected and reviewed prior follow-up interviews conducted face to face and one on one with each student.

Students described the ways in which they considered that the critical elements related to the education experience that they had undertaken and considered whether they believed that they had been appropriate and appropriately applied.

PHASE 4: Presentation of findings

The fourth phase of this research plan was that of determining and presenting the final set of critical elements to be placed into the revised framework. At the completion of this process it was possible to consider the ways in which these critical elements could facilitate the design and assessment of complex and authentic tasks.

Description of the principles

Overall, particularly on application to students, it appeared that the elements considered as critical to the determination of an authentic assessment activity, and as used in the design and development of this module had been appropriate.

With reference to the degree to which challenge is considered to be of importance as a determinant, students seemed satisfied that it was both appropriate and of value as a determinant of authentic assessment activity. Though a degree of concern did exist amongst students as to whether the expected outcome of the assessment activity could be split between performance and product when it may be representative of both. Further consideration led to the conclusion that, in the light of this feedback, the available options were either to revise this critical element to make the split between

performance and *product* less explicit, or to decide whether their was sufficient differentiation between the requirement for authenticity in either *performance* or *product*, that it would be worth splitting them into two separate critical elements. A third possibility was to consider as to whether there was a real need to describe an assessment outcome in terms of *product* or *performance* as it might be more suited as a sub-component of determining *transfer* of learning.

The requirement to demonstrate *metacognition* whether it be by means of *critical reflection*, *self-assessment* or *evaluation* was viewed as being critical in the determination of the authenticity of both a learning module as well as its assessment. However, from the responses received, it seemed apparent that it was the role of *metacognition* itself that was considered to be critical irrespective of the means by which it was mediated, that is whether it was by *critical reflection*, *self-assessment* or *evaluation*. Thus the element should be revised to remove these descriptors of the ways in which metacognition may be demonstrated. It was considered that this would remove the concern that by providing descriptors of the methods by which metacognition may be observed it could be misinterpreted as these being considered as the only means by which metacognition may be demonstrated.

On review of the responses received from students it was felt that the assessment to Module 10 had provided insufficient opportunity to apply the completed assessment outcome against a real workplace application. As a result, the element was revised to ensure that a stronger workplace related link was made.

In terms of the requirement for *fidelity* in the *assessment environment* and the *assessment tools* a strong endorsement was indicated for these as critical elements in the

consideration of authenticity within assessment. However, it was noted that, where possible, the tool used for the assessment should be identical to those used in the workplace.

Finally, the importance of discussion and feedback were demonstrated as critical elements in determining authenticity, as was the requirement to ensure provision for collaboration in any truly authentic assessment.

Findings of the study

The principal question for consideration through this research had been that of determining whether it was possible to determine the principles of authentic activity, to design them into an applicable framework and then use this framework to guide the design, development and application of a more meaningful, more authentic, assessment activity and thus seeking to establish whether *authentic assessment* could provide an effective model for task design and assessment.

The findings in relation to this overall question demonstrate that it is possible to establish, from the research, those elements of design that are considered as critical in the development of authentic assessments. It was also shown that it is possible to develop those elements into a framework that can be used in the design of assessments.

For the purposes of this study the framework developed was applied in the design and development a learning module and its assessment activity that were designed to be capable of flexible delivery. The data collected during the conduct of this study demonstrates that, at least in this instance, it was possible to develop the critical elements of authenticity into an assessment activity, delivered within a flexible learning

environment that remained consistent with the principles of authenticity in assessment and assessment outcome.

In fact, even though these conclusions are necessarily based upon a limited data set it can be argued that authenticity, once deconstructed to determine its critical elements, can present as an effective model for task design and assessment within flexible learning environments.

In order to determine the answer to this principal research question, it has, as previously stated, been necessary to address the following subordinate questions:

1. What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?

2. How do students respond to tasks designed to incorporate the characteristics of authentic assessment?

Principal research question

The principal research question under examination in this study is:

To establish whether it is possible to determine the principles of authentic activity, to design them into an applicable framework and then use this framework to guide the design, development and application of a more meaningful, more authentic, assessment activity, thus seeking to establish whether authentic assessment could provide an effective model for task design and assessment?

The findings in relation to this question, demonstrate that it is possible to identify, from the research, those elements of design that are considered as critical in the development of authentic assessments, and then develop those elements into a framework that can be used in the design of assessments that need to demonstrate a higher degree of authenticity.

For the purposes of this study the framework developed was applied in the design and development a learning module and it's subsequent or integrated assessment activity. The data collected during the conduct of this study demonstrates that, at least in this instance, it had been possible to use these critical elements of authenticity in the re-design of an assessment activity, delivered within a flexible learning environment that remained consistent with the principles of authenticity in assessment and assessment outcome.

In fact, it is possible to deduce from the data for collected by this study that authenticity, once deconstructed to determine its critical component elements, can present as an effective model for task design and assessment. Moreover, it is by considering the ways in which the individual elements of authenticity have been addressed within the design and development of a given task or assessment, that a designer of educational outcomes can state in any measured way that a task or assessment is authentic.

Subordinate research question 1

The first of the required subordinate research question: *What are the specific characteristics of authentic assessment that facilitate design and assessment of complex and authentic tasks?* required the determination as to the elements that would need to be included within the design of the framework of critical elements. It was by means of the

initial determination of these from the review of the literature, and their subsequent refinement, firstly, by means of their iterative review during their development process and more formally via expert review, that enabled the initial framework to be developed, and then applied in the design of the Module 10 and its assessment. The process of data collection and review based upon the delivery of that module and its assessment further enabled the development of the list of critical elements that provided the answer to this first subordinate research question.

Based upon this review, these elements came to be represented as a series of eight *critical questions* that an assessment designer should ask of a given assessment activity in order to determine the degree of authenticity that may be considered as being implicit within that design. Each of these questions is based upon a particular principle of authenticity.

Thus the **first** of the elements that focuses upon the importance of determining that a student will be challenged by an assessment activity is represented as: **To what extent does the assessment activity *challenge* the student?** The intention of this question is to acknowledge that challenge is a vital component of any workplace activity and, to enable the assessment designer to consider the relative degree to which they believe that an assessed student would be challenged by the assessment that they are designing.

Consideration is next given to the **second** question: **Is a *performance, or product,* required as a final outcome?** In this question consideration is given as to whether the outcome of the assessment activity itself will be manifested in terms of a tangible and applicable, workplace orientated workplace, or is the requirement simply to *produce* or *perform* for means of the education or training context only. Here the question asked of

the educational designer, in order to address a matter that was demonstrated to have been of concern for the students within the study group. They made it clear that in their opinion the final application of the assessment activity that they were undertaking did reflect upon both the value and the authenticity of the activity.

The **third** question: **Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?** Focuses upon the notion that, to be considered as authentic, an assessment activity must enable an assessed student to be able to demonstrate that they are capable of transferring learning that they have acquired during an earlier phase of the education process, or within other contexts or domains, from one domain to another, and at the same time, combine it with knowledge and skills from other domains, when, and if, required.

The fourth question for consideration by the assessment designer is: **Does the assessment activity require that *metacognition* is demonstrated?** Within this element the designer is asked to reflect upon the degree to which a student, in undertaking the assessment activity, is given an expectation that the ability to successfully complete an authentic activity will often necessarily require the application of critical reflection, self-assessment or evaluation.

The **fifth** question: **Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?** Is a reflection of the practical notion of workplace where the final determination of successful performance will more than likely be determined by a client, either internal in the form of a workplace superior or colleague, or external, in the form of a paying client? In this respect then, successful authentic assessment requires an awareness of the external validation or judgement of the assessment's outcome.

The **sixth** question: **Is *fidelity* required in the assessment environment? And the assessment tools (actual or simulated)?** Requires that the educational designer is asked to reflect more upon the degree to which the environment within which the assessment is being undertaken may be considered to be authentic to that which the activity would be performed within the workplace. Thus, even though the outcome of an assessment activity could well be the development of a product to be directly applied within the workplace, it is possible that the fidelity of the environment within which it was developed was low. This means that successful performance of the outcome within the assessment environment may not of itself necessarily be reflective of an ability to perform as successfully within the work environment within which a range of additional factor may come into play. It is not to underestimate the value of a positive outcome in a low fidelity environment, more a recognition that it such should not be assumed either.

In consideration of fidelity as a matter of importance in determining authenticity, next is the consideration of the tools used within the assessment environment. In this respect is the recognition that fidelity, particularly of tools to be manipulated in the achievement of an outcome, can be effective when they are virtual or simulated in the same way as when they are actual. Therefore, it may not be necessary that any tools used are those as used in the performance environment, but they should have sufficient fidelity to actual tools used in order that the outcomes of a successful performance in a virtual environment can be juxtaposed to a real environment.

The **seventh** question: **Does the assessment activity require *discussion* and *feedback*?** The penultimate question seeks to ensure that the ability for the student to discuss with one another and, where able, receive feedback, is viewed as both necessary and implicit

to the process of producing the outcome, and not simply viewed as something that happens if time allows. In this respect, discussion and giving and receiving feedback upon the work underway is seen as being consistent with high level performance within a work environment and, therefore, should be considered as necessary to reflecting authentic performance at assessment.

The final and **eighth** question: **Does the assessment activity require that students *collaborate*?** Reflects the important role that collaborating with peers plays within successful workplace performance, and yet, it is a factor that is not always considered within assessment design, where students are often expected to produce individual outcomes alone and without the benefit of collaboration, a circumstance that is often not consistent with authentic workplace interaction and expectation.

Subordinate research question 2

To answer the second of the research questions: *How do students respond to tasks designed to incorporate the characteristics of authentic assessment?* It is necessary to reflect back upon the responses provided by the students, both noted by observation and reported by questionnaire and interview during and immediately after the completion of the module. In this respect it seems that overall students, at least in this instance, considered the use of the principles of *authenticity* to guide assessment design as both appropriate and successful.

Whilst they did raise a number of issues pertinent to the overall design of the module, not the least of which being concerns with the extent to which it had encouraged *collaboration* where they perceived the assessment activity as 'highly individualised'. In addition, concerns had been raised by them about the ways in which they had not

been provided with sufficient access to a template that set out the way that the Defence Force, in particular Army currently met its educational multimedia evaluation requirement. However, on balance, it appears that these issues reflected more concerns with the way in which the researcher, as educational designer, had actually applied the elements, as opposed to the elements themselves.

Other concerns raised related to matters such as the timeframe within which the activity had been undertaken as well as the comparative degree of inexperience that they as students had brought into this process. Others views had related to the degree to which the evaluation tool that they had developed had been applied in the evaluation of real workplace educational multimedia packages. Again, as with the earlier noted concern, this was more a factor of application of the elements as opposed to a commentary on the propriety of the elements themselves. In fact the desire of the students to utilise the evaluation tools that they were developing in a workplace relevant context suggests, to some extent anyway, that they were very much engaged in the *authenticity* of this learning and assessment experience, and in fact wished to extend *authenticity* even further.

In conclusion, it would appear from the student's responses that they students had responded well to a task that had been designed to incorporate the characteristics of authentic assessment. Of particular note being their clear understanding of the ultimate workplace benefits of having to produce authentic outcomes within authentic environments with the use of authentic tools.

Conclusion

Set out below, are the specific characteristics of authentic assessment that, in the considered opinion of this researcher, are critical in the design of authentic assessment tasks. Reflection upon each of these elements, expressed in the form of a question, in the design and development of an assessment activity should assist the educational designer to more effectively ensure that they have better applied the principles of authentic assessment in the design of more authentic and hopefully more valid and reliable assessment activities.

1. To what extent does the assessment activity *challenge* the student?

2. Is a *performance, or product*, required as a final assessment outcome?

3. Does the assessment activity require that *transfer* of learning has occurred, by means of demonstration of skill?

4. Does the assessment activity require that *metacognition*, is demonstrated?

5. Does the assessment require a product or performance that could be recognised as authentic by a client or stakeholder?

6. Is *fidelity* required in the assessment environment? And the assessment tools (actual or simulated)?

7. Does the assessment activity require *discussion* and *feedback*?

8. Does the assessment activity require that students *collaborate*?

Limitations of the study

The findings of this study are supportive of the view that the elements that determine authenticity can be compiled into a framework. Furthermore, they suggest that this framework can be applied in the design of assessment activities that may be acknowledged as being authentic, and that students undertaking such activities are both aware of, and supportive, of the value of authenticity as a determinant in assisting them to acquire and develop the skills and knowledge that they will need to perform successfully within the workplace.

However, in acknowledging these outcomes, it is important that certain limitations within the design, construct and delivery of this study are recognised. The first of these is to acknowledge that this activity was conducted within the Australian Army, by a researcher who was a serving Army officer at the time and using a group of students that were also serving members of the Australian Defence Force. This means that the researcher and the students were, as professional Army training staff, well-used to operating within a competency-based training environment with a significant focus upon the design and delivery of vocational learning outcomes. In addition, each of the students undertaking and reporting back upon this activity were adults, so whilst care would have to be taken in generalising any conclusions to more academically orientated educational outcomes, care would also have to be taken in applying the conclusions of this study to the design of assessment activities for younger students.

An additional limitation was that due to time constraints the researcher had no opportunity to further explore why it was that, during the conduct of the classroom activity, students elected not to collaborate with one another on the development of a

shared assessment outcome, and instead, each set out to develop their own individual outcome. It is possible that this was a function of the design of the assessment activity in that the students were informed that the completed assessment outcome would be a tool that they could take with them and apply, in their subsequent work environment, though it is noted that they could equally have done this with a tool designed and developed collaboratively. It might also have been a function of the military training culture within which these students operated where, particularly in the individual, as opposed to the collective, training environment students were ultimately assessed on individual performance.

Recommendations for further research

As this research activity was undertaken within a design-based research framework the requirement for subsequent research is an acknowledged component of this iterative process. In this regard the first recommendation for the further research in this area is the implementation of the revised framework in the re-design and delivery of this training package within the next two years. In particular, it is recommended that when this re-designed framework is applied particular attention be paid to the evaluated outcomes of those elements or questions that were not revised on this occasion based upon insufficient student feedback, that is the third, fifth, sixth, seventh and eighth critical questions, against each of which students provided feedback that merited more detailed review and reflection in a subsequent application of these elements. It is also recommended that consistent with the original re-design intent of this course to design it to be capable of distance delivery, that it be re-trialled in distance delivery mode.

In order to assist future research it is recommended that the critical elements framework itself be developed into a heuristic for application by educational designers in the assessment design process, a means of enabling them to more formally consider authenticity as a factor in good assessment design, and at the same time, to provide them with a means of measuring the degree to which they has been able to apply authenticity into their assessment designs, and where they were found to be insufficiently authentic, to be able to determine within which of the critical questions they needed to more clearly demonstrate authenticity. Such a framework could also be used in the design and development of a web-based heuristic tool that could assist in future implementation.

Finally, it might be of subsequent research value to seek a means of applying the critical questions individually as a means of seeking their value relative to one another, that is, to establish whether their might be an applicable order of priority in the application of these elements.

REFERENCES

Albion, P. (1999). Heuristic evaluation of educational multimedia: From theory to practice. In J. Winn (Ed.), *Proceedings of the 16th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 9–15). Brisbane: ASCILITE.

Alexander, P.A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, *61*(3), 315–343.

Alexander, S. & Hedberg, J.G. (1994). Interactive multimedia in university education: designing for change in teaching and learning: Evaluating technology-based learning: Which model? *Proceedings of the IFIP TC3/WG3.2 Working Conference on the Design, Implementation and Evaluation of Interactive Multimedia in University Settings* (pp. 233–244). New York: Elsevier.

Anderson, A., McAteer, E., Tolmie, A., & Demissie, A. (1999). The effect of software type on the quality of talk. *Journal of Computer Assisted Learning, 15*(1), 28–40.

Archbald, D. & Newmann, F.M. (1998). *Beyond standardized tests: Assessing authentic achievement in the secondary school*. Reston, Va: National Association of Secondary Principals.

Aurebach, E. (2002). *About alternative assessment*. Auerbach & Associates, Inc. Retrieved May 24, 2004, from http://www.aurbach.com/alt_assess.html

Australian Defence Force. (2007). *The defence training model*. Australian Defence Publication 7.0.2. Canberra: Defence Publishing Service

Avellis, G., & Finkelstein, A. (2002). How to annotate educational multimedia with non-functional requirements. *Educational Technology and Society*, *l? 5*(2), 119–124.

Avery, P.G. (2000). Authentic student performance, assessment tasks, and instruction [Electronic version]. *Research/Practise, 18*(1), 2–12.

Babchuk, W. A. (1996). Glaser or Strauss?: Grounded theory and adult education. *Proceedings of Midwest Research-To-Practice Conference in Adult, Continuing and Community Education.* (pp. 35–51). East Lansing: Michigan State University.

Barzun, J., & Graff, H. F. (1992). *The modern researcher.* (5th ed.). Orlando: Harcourt Brace Jovanovich.

Bayer, A. (1990*). Collaborative-apprenticeship learning: Language and thinking across the curriculum, K-12*. Mountain View, CA: Mayfield.

Benazet, P. (2001). Multimedia evaluation by semiotic approach [Electronic version]. *Journal of International Forum of Educational Technologies and Society and IEEE Learning Technologies Task Force. vol(issue),* 22–26.

Berlak, H. (1992). The need for a new science of assessment. In H. Berlak (Ed.), *Toward a new science of educational testing and assessment*. Albany: State University of New York.

Bigge, M.L. (1976). *Learning theories for teachers.* (3rd ed.). New York: Harper and Row.

Biggs, J. & Tang, C. (2007). *Teaching for quality learning at university. (3rd ed.).* City*:* Open University Press.

Bishop, M.J., & Cates, W. M., (2001). Theoretical foundations for sound's use in multimedia instruction to enhance learning*. Educational Technology Research and Development, 49* (3), 5–22.

Bocij, P., & Greasley, A. (1999). Can computer-based testing achieve quality and efficiency in assessment? [Electronic version]. *International Journal of Educational Technology, 1*(1).

Boud, D. (1995) Assessment and learning: contradictory or complementary?, In P.T. Knight (Ed.), *Assessment for learning in higher education* (pp 35–48). London: Kogan Page.

Brokenbrough, A. S., Chiero, R., & Hoffman, R.P. (1996). *Mapping more authentic learning environments*. Englewool Cliffs, N.J: In Educational Technology Publications,

Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2,* 141–178.

Brown, G.,& Craig, M. (2004). Assessment of authentic learning. Retrieved September 13, 2002, from http://www.coe.missouri.edu/~vlib/glenn.michelle's.stuff/GLEN3MIC.Htm

Brown, J.S., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher,* January–February, 32–42.

Brown, S, Rust, C. & Gibbs, G. (1994). *Strategies for diversifying assessment in higher education.* The Oxford Centre for Staff Development.

Bruffee, K.A. (1984). Collaborative learning and the "conversation of mankind". *College English, 46*(7), 635–652.

Bruner, J.S., (1968). *Toward a theory of instruction.* Toronto: George J. McLeod Ltd.

Burke, K. (1997). *How to assess authentic learning.* Hawker Brownlow Education, Australia.

Chalmers, D. (2007). *A review of Australian and international quality systems and indicators of learning and teaching.* Chippendale, NSW: The Carrick Institute.

Cizek, G.J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practice,* Summer, 16–23.

Collins, A., 1992, Towards a design science of education, In Scanlon, E. and O'Shea, T. (Eds), *New directions in educational technology* (pp. 15–22). Springer, Berlin.

Crooks, T. (1998). Assessing student performance. *HERDSA Greeen Guide No. 8.*

Cummings, J.J. & Maxwell, G.S. (1999). Contextualising authentic assessment. *assessment in education, 6*(2), 177–194.

Custer, R.L. (2000). *Authentic assessment-basic definitions and perspectives.* California Adult Literacy Professional Development Project. Retrieved June 16, 2006, from http://www.calpro-online.org/eric/docs/custer/custer2.pdf

DeCastro-Ambrosetti, D., & Cho, G. (2005). Synergism in learning: a critical reflection of authentic assessment*. The High School Journal, 89*(1), 57–63.

Dede, C. (1995). *The evolution of constructivist learning environments: Immersion in distributed, virtual world*s. Englewool Cliffs, N.J: Educational Technology Publications.

Dixon, Felicia, A. (2000). *The discussion examination: Making assessment match instructional strategy.* Bloomfield Hills: Roeper School.

Dowsing, R.D., & Long, S. (1997). The do's and don'ts of computerising IT Skills assessment. *Proceedings of the 14th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Perth: ASCILITE.

Driscoll, M. P., & Dick, W. (1999). New research paradigms in instructional technology: An Inquiry. *Educational Technology, Research and Development 47*(2), 7–18.

Duffy, T. M., & Cunningham, D. J. (Eds.). (1996). *Constructivism: Implications for the design and delivery of instruction.* New York: Simon & Schuster Macmillan.

English, L. M., (2000). Into the 21st century with spirit. *New Horizons in Adult Education, 14*(1), 1–8.

Ertmer, P.A. & Newby, T.J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 6*(4), 50–72.

Estes, F., & Clark, R.E. (1999). Authentic educational technology: The lynchpin between theory and practice. *Educational Technology, 39(6),* 5–13.

Frohlich, R. (1998). New media communication technologies for facilitating asynchronous delivery of distance learning for differing learning styles: Affective pedagogical techniques for multimedia into the new millennium. *Proceedings of the 15th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Wollongong: ASCILITE.

Gagné, R. M., Briggs, L. J., & Wager, W.W. (1992). *Principles of instructional design.* (4th ed) Fort Worth: Harcourt, Brace Jovanovich.

Gay, L.R. (1996). *Educational research: Competencies for analysis and application* (5th ed). Florida: Prentice-Hall.

Gazzard, S. & Dalziel J.R. (1998). Comdesign principles for the 'next wave' educational tools: The development of the WebMCQ assessment system. *Proceedings of the 15th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Wollongong: ASCILITE.

Gibran, K. (1992). *The prophet.* Penguin Books.

Glaser, B.G. (1965). The constant comparative method of qualitative analysis. *Social Problems, 12* (4), 436–445.

Glaser, B.G. & Anselm, S.L. (1967). *The discovery of grounded theory: strategies for qualitative research.* Aldine Publishing.

Haney, W. & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. Phi Delta Kappan, 70 (9), (pp 683–687).

Harper, B., & Hedberg, J. (1997). Creating motivating interactive learning environments: A constructivist view. *Proceedings of the 14th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Perth: ASCILITE.

Hart, D. (1994). Authentic assessment: A handbook for educators. California: Addison-Wesley Publishing Company, Inc.

Hattie, J., Biggs, J. & Purdie, N. (1996, Summer). Effects of learning skills interventions on student learning: A meta-analysis. Review of Educational Research, 66(2), 99–136.

Herman, J.L., Aschbacher, R. & Winters, L. (1992). *A practical guide to alternative assessment.* Alexandria, VA: ASCD publications.

Herrington, A., & Moran, C. (2001). What happens when machines read our students' writing?' *Proceedings of National Council of Teachers of English Conference on College Composition and Communication.* Champagne, Illinois: NCTE.

Herrington, J., & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Education Research and Development, 17*(3), 305–322.

Herrington, J. & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology, Research and Development, 48*(3), 23–26.

Herrington, J., Oliver, R., Herrington, T., & Sparrow, L. (1997). Enhancing transfer from interactive multimedia to real-world practice. *Proceedings of the 14th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Perth: ASCILITE.

Herrington, J., Reeves, T.C. & Oliver, R. (2006). Authentic tasks online: A synergy among learner, task, and technology. *Distance Education, 27*(2), 233–247.

Herrington, J. & Standen, P. (1997). Acumen: An interactive multimedia simulation based on situated learning theory. *Proceedings of the 14th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Perth: ASCILITE.

Herrington, J & Standen, P. (2000). Moving from an instructivist to a constructivist multimedia learning environment. *Journal of Multimedia and Hypermedia*, *9*(3), 195–205.

Hoepfl, M. (2000). Large-scale authentic assessment. In Custer, R. (Ed.), *Authentic Assessment in Vocational Education,* pp. 49–67. Columbus, OH: ERIC/ACVE

Jeffries, P., McDaniel, R. & Vaughn, M. (1998). Development of an interactive, multimedia CD ROM to teach medication administration to undergraduate health professionals. Indiana Higher Education Faculty. Retrieved July 16, 2002, from http://www.ihets.org/learntech/distance_ed/fdpapers/1998.

Jonassen, D.H. (2001). (Ed.) *Handbook of research for educational communications and technology.* New York: Macmillan.

Jonassen, D. H. (1994). Thinking technology toward a constructivist design model. *Educational Technology*, April, pp 34–37.

Jonassen, D.H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology Research and Development*, 39(3), 5–14.

Jonassen, D.H., & Hernandez-Serrano, J. (2002) Case-based reasoning and instructional design: Using stories to support problem solving. *Educational Technology Research and Development, 50*(2), 65–77.

Jonassen, D.H. & Kwon, H.I. (2001). Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development, 49*(1), 31–51.

Jonassen, D.H., Tessmer, M. & Hannum, W.H. (1999). *Task analysis methods for instructional design.* Mahwah, N.J: Lawrence Erlbaum Associates.

Jones, D. (1997). Three generations of online assignment management. *Proceedings of the 14th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Perth: ASCILITE.

Khattri, N., Reeve, A.L. & Kane, M.B. (1998). Principles and practices of performance assessment. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Kellehear, A. (1993). *The unobtrusive researcher: A guide to methods.* St. Leonards, NSW: Allen & Unwin.

Kendle, A. & Northcote, M. (2000). The struggle for balance in the use of quantitative and qualitative online assessment tasks. *Proceedings of the 17th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Coffs Harbour: ASCILITE.

Kendle, A. & Northcote, M. (2001). Online assessment criteria in action: Task design in contrasting tertiary education contexts. *Proceedings of ED-MEDIA World Conference on Educational Multimedia, Hypermedia & Telecommunications.* (pp. 920–940). Seattle, WA: Association for the Advancement of computing in Education.

Kerka, S. (1995). Techniques for authentic assessment practice application brief. Center on Education and Training for Employment. Retrieved July 15, 2002, from http://www.cete.org/acve/docgen.asp?tbl=archive&ID=A032.

Khattri, N., Reeve, A.L. & Kane, M.B. (1998). *Principles and practices of performance assessment.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Laurillard, D., Stratfold, M., Luckin, R., Plowman, L. & Taylor, J. (2000). Affordances for learning in a non-linear narrative medium. *Journal of Interactive Media in Education, 2*.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge: Cambridge University Press.

Lonsdale, M., & McCurry, D. (2004). Literacy in the new Millenium. Report for the Australian Council for Educational Research. Retrieved November 30, 2006, from http://www.ncver.edu.au/research/proj/nr2L02.pdf?PHPSESSID=9e837f133741ad7aaffe110eb42627fb

Lowe, John, P. (1994). Assessment that promotes learning. Schreyer Institute for Teaching Excellence Pennsylvania State University. Retrieved March 25, 2003, from http://www.schreyerinstitute.psu.edu/Resources/Assessment-Learning.asp.

Lund, J. (1997). Authentic assessment: Its development and applications. *Journal of Physical Education,Recreation, and Dance,* 68(7), 25–28.

Martens, R. & Hermans H. (2000). Internet based formative prior knowledge assessment. *Studies in Educational Evaluation, 26,* 245–258.

Mabry, L. (1999). *Portfolios plus: A critical guide to alternative assessment.* Thousand Oaks, CA: Corwin.

McAlister, B. (2000). The authenticity of authentic assessment: What the research says…Or doesn't say. University of Wisconsin-Stout. Retrieved January 12, 2005, from.http:ericacve.org/textonly/mp_mcalister_03.asp

McCormick, J. (2000, February). On a high-tech firing line. *Newsweek, 12–15*.

McLellan, H. (1994). Situated learning: Continuing the conversation. *Educational Technology*, 34 (10), pp. 7–8.

McLoughlin, C. & Luca, J. (2000). Assessment methodologies in transition: Changing practices in web-based learning. *Proceedings of ASET Conference.* Toowoomba: University of Southern Queensland.

McMillan, J.H. (2000). Fundamental assessment principles for teachers and school administrators [Electronic version]. *Practical Assessment, Research & Evaluation, 7*(8).

Merriam, S.B. (1998). *Qualitative research and case study applications in education.* San Francisco: Jossey-Bass.

Merson, J. (2001, May 12). How the mind remembers. *Sydney Morning Herald [Good Weekend Magazine], 5–9.*

Moon, T.R. & Callahan, C.M. (2001). Classroom performance assessment: What should it look like in a standards-based classroom? *National Association of Secondary School Principals. NASSP Bulletin, 85*(622), 48–58.

Moorcroft, T.A., Desmarais, K.H., Hogan, K. & Berkowitz, A.R. (2000). Authentic assessment in the informal setiing: How it can work for you. *Journal of Environmental Education, 31*(3), 20–25.

Nadolski, R.J. (2001). A model for optimizing step size of learning tasks in competency-based multimedia practicals. *Educational Technology Research and Development, 49*(3), 87–101.

Neimeyer, G.J. (1993). Constructivist assessment: A casebook. Newbury Park, CA: Sage.

Newmann, F.M., Marks, H.M. & Gamorran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education, 104*(4), 280–312.

Newmann, F.M., Lopez, G. & Bryk, A.S. (1998). The quality of intellectual work in Chicago schools: A baseline report. Chicago: *Consortium of Chicago School Research.*

Newmann, F.M. & Wehlage, G.G. (1993). Five standards of authentic instruction. *Educational Leadership*, *50*(7), 8–12.

Perkins, D.N. (1991). Technology meets constructivism: Do they make a marriage? *Educational Technology, 31*(5), 18–23.

Pham, B. (1998). Quality evaluation of educational multimedia systems. *Australian Journal of Educational Technology, 14*(2), 107–121.

Prestidge, L.K. & Glaser, Williams, C.H.W. (2000) Authentic assessment: Employing appropriate tools for evaluating students' work in 21st century classrooms. [Electronic version]. *Intervention in School and Clinic.* Jan 2000.

Puckett, Margaret B. & Black, Janet K. (1994). *Authentic assessment of the young child — Celebrating development and learning.* New Jersey: Prentice-Hall.

Pyne D. (1994). *Designing educational projects and program evaluation.* Boston: Kluwer Academic Publishers.

Ramsden, P. (2004). *Learning to teach in higher education* (2nd ed.). New York: Routledge, Falmer, Taylor and Francis Group.

Reeves, T.C. (2006). Design research from a technology perspective. In Van den Akker, J., Gravemeijer, K., McKenney, S. & Nieveen, N. (Eds). (2006). *Educational Design Research*. London: Routledge, Taylor and Francis.

Reeves, T.C. (2000). Socially responsible educational technology research. *Journal of Educational Technology*, *4*(6), 19–28.

Reeves, T.C. (1992). Evaluating interactive multimedia. *Educational Technology, 32*(5), 47–53.

Reeves, T.C. (1992). Effective dimensions of interactive learning systems. *Proceedings of the Information Technology for Training and Education — ITTE Conference.* (pp. 99–115). Brisbane: ITTE.

Reeves, T.C., Laffey, J.M. & Marlino, M.R. (1991). Using technology as cognitive tools: Research and praxis. Proceedings of the 14th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education. Perth: ASCILITE.At *ASCILITE '97*, Perth, WA. *ASCILITE '97*, Perth, WA.

Reeves, T.C. & Okey, J.R. (1996). *Alternative assessments for constructivist learning environments.* Engelwood Cliffs, N.J: Educational Technology Publications.

Reigeluth, C.M. (1991). Reflections on the implications of constructivism for educational technology. *Educational Technology*, 31 (9), 34–37.

Reiser, R.A., (2001). A history of instructional design and technology: Part 1: A history of instructional media. *Educational Technology Research and Development, 49*(1), 53–64.

Rossman. M. (2000). Andragogy and distance education: Together in the new millennium. *New Horizons in Adult Education, 14*(1).

Rowntree, D. (2000). *Assessing students: How shall we know them?* London: Nichols Publishing.

Rudner, L.M. and Boston, C. (1994). 'Performance assessment', ERIC Review, 3 (1), 2–12.

Sadler, R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education, (13)*2, 191–209.

Schank, R. (1994). *Engines for educators.* Retrieved January 12, 2005, from http://www.engines4ed.org/hyperbook/nodes/educator-outline.html

Schell, J.W. (2000). *Think about authentic learning and 'then' authentic assessment.* Retrieved January 12, 2005, from http:ericacve.org/textonly/mp_schell_02.asp

Scott, J. (2000). Authentic assessment tools. In R. L. Custer (Ed.), J. W. Schell, B. McAlister, J. Scott, & M. Hoepfl. *Using authentic assessment in vocational education.* Information Series No. 381 (pp. 40–55).

Shellnut, B., Knowlton, A, & Savage, T. (1999). Applying the ARCS model to the design and development of computer-based modules for manufacturing engineering courses. *Educational Technology, Research and Development, 47*(2), 00–110.

Shironica K., Ferry, B. & Hedberg, J. (2001). *Developing internet-based study materials as novice web-designers: Experiences of twelve Sri-Lankan teacher educators*. *Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.* Melbourne: ASCILITE.

Shute, V.J. and Psotka, J. (1996) In *Intelligent tutoring systems: Past, present, and future.* In Jonassen, D.H. (ed). Handbook of Research for Educational Communications and Technology.

Squires, D. (1997). An heuristic approach to the evaluation of educational multimedia software. *Proceedings of the CAL 97 Conference*. Exeter: University of Exeter.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation.* Washington, DC: Brookings Institution Press.

Tanner, D.E. (1997). The long (suit) and the short (comings) of authentic assessment. Paper presented at the *Annual Meeting of the American Association of Colleges for Teacher Education* (49th, Phoenix, AZ, February 26–March 1, 1997)

Tasker, R., (1996). The Vischem Project: Using multimedia to develop mental models of the invisible molecular world. *Proceedings of the AUC Academic Conference — From Virtual to Reality*. Brisbane: The University of Queensland.

The Carrick Institute. (2006). *Strategic directions for 2006.* Canberra: The Carrick Institute of Learning and Teaching in Higher Education.

Tuckman, B.W. (1978). *Conducting educational research (*2nd ed.). New York: Harcourt Brace Jovanovich

Van den Akker, Jan, Gravemeijer, Koeno, McKenney, Susan and Nieveen, Nienke. 2006 (Eds). *Educational Design Research*. Published by Routledge, Taylor and Francis Group, London and New York

Van den Akker, J. (1999). *Principles and methods of development research.* In J.van den Akker, N. Nieveen, R.M. Branch, K.L. Gustafson, & T. Plomp, (Eds.), *Design methodology and developmental research in education and training (pp. 1–14).* The Netherlands: Kluwer Academic Publishers.

Van Merrienboer, Jeroen J.G., Clark, Richard E., & de Croock, Marcel B.M. (2002). *Blueprints for complex learning: The 4C/ID model.* In *Educational Technology Research and Development.* Vol. 50, No. 2, pp 39–63.

Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *The Modern Language Journal, 81*(4), 470–481.

Wenger, E. (2002). Themes and ideas: Communities of practice. Retrieved January 14, 2005, from http:www.ewenger.com/ewthemes.html

Wiggins, G (1993). *Assessing student performance — Exploring the purpose and limits of testing.* San Francisco: Jossey-Bass.

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation, 2*(2). Retrieved September 5, 2006 from http://PAREonline.net/getvn.asp?v=2&n=2

Willyerd, K.A. (1997, March). Balancing your evaluation act. *Training,* 52–58.

Wilson, B.G. (1996). (Ed.). *Constructivist learning environments: case studies in instructional design.* Englewood Cliffs, N.J.: Educational Technology Publications.

Wolf, D., LeMahieu, P., & Eresh, J. (1992, May). Good measure: Assessment as a tool for education reform. *Educational Leadership*, 49(8), 8–13.

APPENDIX 1

EXPERT REVIEWER INTERVIEW QUESTIONNAIRE

Referenced: p. 50

Expert Reviewer Interview Questionnaire

Respondent Name:

Date of Interview:

Interview Questions

1. Do the critical elements make sense?

2. Does each individual element, as set out, cover what may be considered to be critical to the determination of 'authenticity' within an assessment? (*Seek feedback from the respondent on each of the critical elements. Seek to establish what, if any, specific elements, or parts of an element, the interviewee considers has not been covered in each.*)

   - Degree of challenge(s) presented to the assessed student.

     Comment:

   - Performance or product as final assessment outcome.

     Comment:

   - Transfer of learning (skills/knowledge/attitude) required.

     Comment:

   - Critical reflection & self-assessment or evaluation required.

     Comment:

- Accuracy in product or performance, and fidelity of assessment environment, is displayed.

  Comment:

- Fidelity of assessment tools used.

  Comment:

- Discussion and feedback required.

  Comment:

- Collaboration required.

  Comment:

3. Can you provide any feedback or detail on any of the elements that would enhance their suitability or applicability?

- Degree of challenge(s) presented to the assessed student.

  Comment:

- Performance or product as final assessment outcome.

  Comment:

- Transfer of learning (skills/knowledge/attitude) required.

  Comment:

- Critical reflection & self-assessment or evaluation required.

  Comment:

- Accuracy in product or performance, and fidelity of assessment environment, is displayed.

  Comment:

- Fidelity of assessment tools used.

  Comment:

- Discussion and feedback required.

  Comment:

- Collaboration required.

  Comment:

4. Are there any further elements or areas, currently not included, that you consider should be added to those already included?

5. Other issues or comments. (*This should include…)*

APPENDIX 2

STUDENT EVALUATION QUESTIONNAIRE

Referenced: ps. 60, 146 and 183

Student Evaluation Questionnaire

Student Name:

Evaluation Questions

1. How well do you believe each of the critical elements set out below was addressed within Module 10 of the CBL Practitioners Course — 'Evaluate Educational Multimedia'?

   - Degree of challenge(s) presented to the assessed student.

     Comment:

   - Performance or product as final assessment outcome.

     Comment:

   - Transfer of learning (skills/knowledge/attitude) required.

     Comment:

   - Critical reflection & self-assessment or evaluation required.

     Comment:

   - Accuracy in product or performance, and fidelity of assessment environment, is displayed.

     Comment:

- Fidelity of assessment tools used.

  Comment:

- Discussion and feedback required.

  Comment:

- Collaboration required.

  Comment:

2. Can you provide any feedback or detail on any of the elements that you think would have enhanced their applicability?

   - Degree of challenge(s) presented to the assessed student.

     Comment:

   - Performance or product as final assessment outcome.

     Comment:

   - Transfer of learning (skills/knowledge/attitude) required.

     Comment:

- Critical reflection & self-assessment or evaluation required.

  Comment:

- Accuracy in product or performance, and fidelity of assessment environment, is displayed.

  Comment:

- Fidelity of assessment tools used.

  Comment:

- Discussion and feedback required.

  Comment:

- Collaboration required.

  Comment:

3. Are there any further elements or areas that were not included, that you consider should have been included?

4. Additional issues or comments.

APPENDIX 3

STUDENT INTERVIEW QUESTIONNAIRE

Referenced: ps. 61, 146 and 183

Student Interview Questionnaire

Student Name:

Date of Interview:

Interview Questions

After providing a brief outline of the purpose of the research, and based upon the written feedback provided by the student ask the following questions:

1. What do you understand by authenticity in assessment?

2. How well do you believe each of the critical elements set out below was addressed within Module 10 of the CBL Practitioners Course — 'Evaluate Educational Multimedia'? (*Seek feedback from the respondent on each of the critical elements. Seek to establish what, if any, specific elements, or parts of an element, the interviewee considers has not been covered in each module.*)

   - Degree of challenge(s) presented to the assessed student.

     Comment:

   - Performance or product as final assessment outcome.

     Comment:

- Transfer of learning (skills/knowledge/attitude) required.

  Comment:

- Critical reflection & self-assessment or evaluation required.

  Comment:

- Accuracy in product or performance, and fidelity of assessment environment, is displayed.

  Comment:

- Fidelity of assessment tools used.

  Comment:

- Discussion and feedback required.

  Comment:

- Collaboration required.

  Comment:

3. Can you provide any feedback or detail on any of the elements that would have enhanced their applicability?

- Degree of challenge(s) presented to the assessed student.

  Comment:

- Performance or product as final assessment outcome.

  Comment:

- Transfer of learning (skills/knowledge/attitude) required.

  Comment:

- Critical reflection & self-assessment or evaluation required.

  Comment:

- Accuracy in product or performance, and fidelity of assessment environment, is displayed.

  Comment:

- Fidelity of assessment tools used.

  Comment:

- Discussion and feedback required.

  Comment:

- Collaboration required.

  Comment:

4. Are there any further elements or areas that were not included, that you consider should have been included?

5. Additional issues or comments.

APPENDIX 4

COMPUTER BASED LEARNING PRACTITIONERS COURSE —

MODULE 10 — EVALUATING EDUCATIONAL MULTIMEDIA

Referenced: p. 96

Computer Based Learning Practitioners Course

Module 10 — Evaluating Educational Multimedia

Module Content

Module Aim

The aim of this module is to provide trainees with an understanding of the general principles and practices of the evaluation of educational multimedia.

Module Purpose

The module is designed to provide trainees with an understanding of the general principles and practices of evaluation of educational multimedia and its application to the evaluation of Army computer based learning packages.

> '…multimedia should be designed to support the principles that learning involves knowledge construction where new knowledge is built upon existing knowledge and within meaningful contexts.' (Reeves, 1992 in Binh Pham, 1998)

1.0 Explain educational multimedia

'A resource is a multimedia one when it calls simultaneously upon different sensory registers and if it generates interactivity between the learner and the artefact, made up in the majority of the cases of a data-processing device. The interactivity is thus fundamental and constitutes the specificity which distinguishes multimedia from audio-visual'. (Patrick Benazet, 2001)

Introduction

1.1 Educational Multimedia

From the early to mid 1980's, particularly with the advent of personal computers with graphical users interfaces (GUI) capability, trainers and educators began to seek increasingly sophisticated methods of employing these data-processing machines in the delivery and management of education and training packages.

However, crucial to the continual enhancement of the design, development and delivery of training and education, using this medium, is the ability to review and critically analyse these packages as learning resources. Newly-trained classroom teachers and instructors must learn to review the quality of teaching of themselves and others, in the same way Computer Based Learning Practitioners need to be able to critically reflect upon the educational multimedia content which they use, observe and design. Good educational multimedia must maintain its focus on the educational requirement, while at the same time ensuring that it makes the most of what technology has to offer.

1.2 Main approaches in the delivery of educational multimedia

Prior to evaluating an educational resource some understanding has to be achieved as to the theoretical educational approach employed by the designer.

There have been many attempts to classify the ways in which people learn. From these a range of theories has been evolved, each of which provides a methodology or blueprint to be used in the design of learning content. Set out below, are three of the major

theories most usually represented in educational design, learning outcomes, cognitive and affective or psychomotor.

Learning Outcome

One of the most well known theorists on the ways in which people learn, and, therefore, the ways that instruction should be designed, was Gagné. He believed that it is possible to identify learning outcomes in a body of knowledge or skill. These learning outcomes can then be systematically identified and measured in terms of knowledge by performing tasks and sub-tasks, which are organised in a hierarchical fashion (Gagné and Briggs, 1997, in Binh Pham (1998)).

Cognitive

The cognitive approach believes that knowledge acquisition is cumulative and not necessarily hierarchical. This view shifts the instructional design emphasis from teaching to learning, moreover if knowledge acquisition is cumulative, as opposed to hierarchical, then the instructional designer of educational multimedia is more able to employ the hyper-linking available within multimedia.

Affective

The affective and psychomotor approach to instructional design places greater focus on establishing the emotional and psychological aspects of the learner's responses, as it is these that will determine the learner's motivation to learn. The motivational theory of Keller and his ARCS Model is an example of this. The model is based on a learner's Attention, the Relevance of the content to the learner, the level of Confidence with which the learner will approach this content and the degree of Satisfaction that the learner will acquire from completion of the learning.

1.3 Types of evidence to be gathered

In order to evaluate educational multimedia, a range of questions must be evolved. In order to determine these questions consideration needs to be given to the types of evidence available to the evaluator, and the individual features of the CBLP that are to be evaluated.

According to Pham (1998), the main areas to be evaluated include:

- The objectives and content of the CBLP
- The quality of the interactivity
- The overall attraction of the CBLP
- The teaching strategy employed within the CBLP
- The data processing reliability of the CBLP

In order to evaluate these areas, consideration must be given to the following questions:

Interface

- How well can trainees apply the interface to the required tasks?
- Does the trainee gain a degree of 'user satisfaction' in using this interface?
- Can the trainee learn to use and operate the system with ease?
- Does the interface encourage effective and efficient performance of specific tasks?

Navigation

- Can the trainee easily obtain knowledge or perform tasks following the links provided?
- Does the information contained in the CBLP facilitate relational understanding of concepts?
- In what ways does the navigation method employed enhance the trainee's ability to learn when compared with a more traditional approach to instruction?
- Does the navigation both enhance creative ideas and encourage a higher degree of trainee commitment?

Assessment Criteria — Formative assessment/quiz

1.1 Define educational multimedia

1.2 Outline the main educational multimedia models

1.3 Describe types of evidence to be gathered forms the evaluation of educational multimedia

2.0 Structure of educational multimedia reports

According to Alexander and Hedberg (1994, in Matshediso (2000)) there have since the 1940's been four broad historical orientations in the evaluation of educational multimedia. However, all of the evaluation of educational multimedia undertaken falls into two broad categories, which can be defined as 'formative' and 'summative' evaluation.

Formative

Formative evaluation is taken as being that which occurs during the multimedia design and development phases. The main purpose of this type of evaluation is to view the production process as an iterative one, with a requirement for an on-going evaluation process to ensure that the package under development continues to be relevant and appropriate to its educational goals.

Summative

Summative evaluation occurs at the completion of the project. It may take place either before hand-over to the client, as a final educational quality assurance check, or subsequent to package hand-over, as a means of reviewing the product delivered and determining what lessons may be learnt from it for future application.

2.1 Approaches and methods for evaluating elements

1940's — Objective-based — Summative

In the 1940's evaluation of educational multimedia resources, albeit not CBLPs at this time, was predominantly objective-based. The intention of this method of evaluation was to determine how successfully the educational objective had been met through the application of multimedia in the training situation.

1970's — Decision-based — Formative

During the 1970's, educational multimedia evaluation became decision-based. This methodology sought to establish a process of evaluating the multimedia at all stages during its development process. The adoption of this 'decision-based' approach meant

that evaluation could become a part of the design process, and not something that occurred subsequent to it.

1980's — Naturalistic — Formative

From the 1980's the evaluation of educational multimedia became more naturalistic in approach. The focus was shifted to the achievement of the educational goals to be achieved and a consideration as to whether, in educational terms, the goals were worth achieving.

1990's and on — Holistic/integrated — Formative/Summative

From the middle of the 1990's onwards, the focus of evaluation of educational multimedia became both more holistic and more integrated. In many respects, it started to be recognised that the evaluation of educational multimedia had a role to play in both the formative, and design and development phases, of multimedia production as well as in the summative, or testing phase. From this stage on a clear trend began to emerge with the evaluators beginning to become more systematic in the evaluation methodology that they applied, with the formulation of clearly defined questions and goals.

2.2 The elements to be evaluated

Three major elements of an educational multimedia system will have a significant impact on its quality. In short these are, the subject matter or content, the ways in which the subject matter is presented and organised and the technical tools used to convey and construct knowledge.

However, prior to commencing the review of a package, the evaluator must seek to establish the overall intention of the package and consider the audience at which it has been aimed. Once this has been completed they may then commence the 'Macro Evaluation' of the package.

'Macro Evaluation'

The macro evaluation is comprised of the following:

- Content is confirmed against doctrine and the TMP
- Clear objectives are stated for the package
- Confirmation is made that these objectives are met throughout the CBLP
- The AST structure of the CBLP is confirmed as being consistent (Intro/Revision/Lesson/Summary/Assessment)
- Information is adequately and logically 'chunked'
- The information appears in either a instructivist or constructivist format depending upon the content requirements
- Confirmatory activities occur at the end of every 'chunk'
- Screens are adequately 'paced' (such as not too much or too little information appearing on screens at any one time)
- Assessment are appropriate, valid and authentic with respect to the content
- The CBLP is technically sound (conforms to Tech Specs, DOMAIN, SCORM, AICC)

- The CBLP is instructionally sound (conforms to TTC's Instructional Design Specs)
- The navigation is intuitive and not cumbersome
- There is a progress indicator
- The user has access to a site map which indicates where they have been and where they have yet to go
- The user has the ability to replay audio, video and animation if required
- All functions work

On completion of the broader components of the macro evaluation of the CBLP consideration must be given to the more detailed factors that make up the micro evaluation.

Micro Evaluation

Text

- Confirm the font and style
- Confirm SPG
- Confirm content accuracy
- Confirm position on screen (people read — left to right, therefore, the text should be positioned on the left of the screen if it is the 'main carrier' of the information. Only the main point need appear as text on screen. Any additional information can appear in other multimedia formats to enhance better understanding of the text message)

- Ensure that the text is not exactly the same as the audio (Split attention)
- Ensure that the user has control of when the text leaves the screen. Each user will read at a different pace, therefore, having control of when the text disappears from the screen is essential. Also, if the CBLP has the 'notebook' function enabled, this allows the user the opportunity to cut and paste the text.
- Ensure the colours of text used is consistent throughout the package (except when using pneumonic), and ensure that the text can be read by users who may have colour perception problems

Audio (Narration)

- Confirm consistency of pitch and levels
- Confirm accuracy of content
- Ensure the audio supports the text (add extra text if required)

Audio (Sound effects)

- Confirm consistency of pitch and levels
- Ensure they are appropriate
- Ensure that they enhance the retention of information (for example Emotion/expectation/suspense)
- Check cost/copyright

Graphics

- Confirm consistency and standardisation of colour, treatment, size, borders
- Ensure that the graphics are appropriate
- Ensure enhances message of the content
- Ensure they are consistent in size and scale with regard to the available screen real estate

Animation/Video

- Confirm consistency and standardisation
- Confirm consistency of pitch and levels
- Ensure appropriate to message of the content (not just for entertainment)
- Ensure add value to the text/audio on screen
- Check cost and copyright

Assessment Criteria — Formative assessment/quiz

2.1 Identify the necessary changes required from evaluation data

2.2 Describe the components of a revision plan

3.0 Apply the process of educational multimedia evaluation to a TTC developed CBLP

Assessment Outline

Trainees will construct and apply their own educational multimedia evaluation tool to a TTC CBLP. On completion, they will, after consultation with at least two other course members, revise their evaluation tool, and then re-apply it to a subsequent TTC CBLP.

This assessment activity will take place in 4 parts.

1. Trainees will construct their own educational multimedia evaluation tool, and apply it to a TTC CBLP.

2. Trainees will then critically review and evaluate the performance of their educational multimedia evaluation tool in the light of feedback from at least two other course members, and produce a (100–250 word) critique. (Note: The review and evaluation process will include peer review of the evaluation tool and consideration of those produced by other course members).

3. The review and evaluation data acquired at 2 above will enable trainees to revise their evaluation tool which they will then apply to another TTC CBLP.

4. Finally, trainees will conclude the assessment with the production of a critical report (250–500 words) of the CBLP noted at 3 above, noting both positive and negative aspects and features, and make recommendations as to this CBLP could be improved.

NOTE:

On completion all the final assessment outcomes are to be forwarded to the OC FDDS-B.