Performance enhancement of the task assessment process through the application of an electronic performance support system

Alistair Campbell
Edith Cowan University

Follow this and additional works at: https://ro.ecu.edu.au/theses

Part of the Educational Assessment, Evaluation, and Research Commons

Recommended Citation

This Thesis is posted at Research Online.
https://ro.ecu.edu.au/theses/18
Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.

- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author’s moral rights contained in Part IX of the Copyright Act 1968 (Cth).

- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
Performance Enhancement of the Task Assessment Process through the Application of an Electronic Performance Support System

Alistair Bruce Campbell
BApplSc, DipEd, PostGradDipEd, MEd, GradDipIMMT

This thesis is presented in fulfilment of the requirements for the degree of Doctor of Philosophy in Education

Faculty of Arts and Education
EDITH COWAN UNIVERSITY

2008
USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.
Not everything that counts can be counted;
and not everything that can be counted counts.

*Albert Einstein*

One certain way to improve the quality of data:
*improve its use!*

*Ken Orr*
ABSTRACT

Higher education in Australia, as for many other countries, has changed greatly over the last 20 years at all levels and in many areas of operation including teaching, learning and assessment. The driving forces for these changes have been both internal and external, and have included factors such as: the increasing student population; the increasing use of part-time staff; a reduction in government funding; an increased expectation of institutional accountability; and the growing access and use of information and communication technologies (ICT) in teaching and learning. Assessment has not escaped these changes but in many cases has not kept up with exemplary and recommended practice. This is especially so in the specific area of task assessment that involves professional judgement, where assessment is a time-consuming, expensive, poorly managed, and a stressful professional activity, and is often a negative emotional experience for both learners and educators.

Professional judgement of educators in the task assessment process is becoming more important with the trend towards student-centred, standards-based curriculum and the use of authentic assessment tasks that are more subjective in nature. At the same time, stakeholders are demanding greater validity, reliability and transparency in the assessment process. To meet these demands, a new re-framing of the task assessment process is required, involving activities such as the design of the marking key, moderation, marking, feedback, reporting and management. Meanwhile, current methods and practices used in the task assessment process that involve professional judgement have not kept pace with current best practice, nor do they involve the application of ICT to any great extent.

The study partly developed from the researcher’s professional reflection on the above issues. The study also developed from the researcher’s observation of the importance that electronic performance support systems (EPSS) and knowledge management (KM) have had in the commercial world in the area of improved performance of the worker and work processes. The premise of the study was to investigate to what extent the performance of the task assessment process involving professional judgement could be improved and enhanced through the application of an EPSS. A preliminary review of the literature identified three fields that needed to be reviewed, investigated and integrated for this study; these were: assessment of achievement, the use of EPSS, and software design and development.
Based on the literature review on the use of EPSS, the human performance technology (HPT) model was selected as the over-arching conceptual framework, and was combined with participatory action research (PAR), participatory design (PD) and co-operative inquiry methodologies to inform and develop the research design. This emergent study used these methods to explore, design, develop, implement and evaluate the study’s premise.

The study found that the developed intervention strategies and the EPSS could significantly improve the performance of learners, tutors and coordinator during the task assessment process that involved professional judgement. The strategies and EPSS augmented, enhanced, integrated and supported the performance of the task assessment process in many ways. The EPSS eliminated or reduced the clerical and administrative or ‘busy’ work that is typically associated the task assessment process. While the intervention strategies that included the development of instructional rubrics, modelled constructivist learning, and in the process facilitated collaborative, practical and educative consequences for students, tutors and unit coordinator. Four assertions were developed based on the study’s findings, two of which related to the improvement and enhancement of the marking key, and two to the improvement and enhancement of the coordinator’s and tutors’ performance.

Confirmative evaluation of the study’s findings has been achieved through the continued use and development of the EPSS over the last four years, as well as through conference presentations, publications and awards, at both university and national level. This has demonstrated that the EPSS and associated strategies can be successfully applied to different types of markings keys and assessment tasks to improve performance of the task assessment.
DECLARATION

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

(i) Incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution or higher education;
(ii) Contains any material previously published or written by another person except where due reference is made in the text; or
(iii) Contain any defamatory material.

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

Signature:

Date: 18 April 2008
ACKNOWLEDGEMENTS

The thesis could not have been completed without the support and encouragement of so many good friends and colleagues.

Firstly, acknowledgements must go to Susan Krieg, whose enthusiasm and willingness to explore the application of technology in the field assessment made this study possible.

Secondly, I acknowledge the development team members: Susan Krieg, Jenny Jay, Sue Sharp, Kerry Williams, Pippa Nelligan, Julia Wren, and Yvonne Haig for their enthusiastic and sustained support and cooperation throughout the Development Phase of the study. In the Post-Development Phase of the project Heather Sparrow, Jenny Lane, John Duff and John Tingay were of great support and encouragement.

Thirdly, I thank my supervisor, Dr Paul Newhouse, for being there for advice and support, and for reading the many drafts that preceded this thesis; and my co-supervisor, Dr Jeremy Pagram.

Finally, I thank my wife Beatriz for travelling this journey with me. We both look forward to the next adventure.
# TABLE OF CONTENTS

**ABSTRACT** .................................................................................................................. vii

**DECLARATION** ........................................................................................................ ix

**ACKNOWLEDGEMENTS** ........................................................................................ xi

**TABLE OF CONTENTS** .......................................................................................... xiii

**LIST OF TABLES** .................................................................................................... xvii

**LIST OF FIGURES** ................................................................................................. xix

**LIST OF ACRONYMS** ............................................................................................. xxi

**GLOSSARY** ............................................................................................................ xxiii

**CHAPTER ONE INTRODUCTION** ............................................................................. 1

- Aim .......................................................................................................................... 3
- Rationale .................................................................................................................. 4
- Significance .............................................................................................................. 7
- Scope of study .......................................................................................................... 9
- Research propositions ............................................................................................. 10
- Thesis overview ...................................................................................................... 11

**CHAPTER TWO LITERATURE REVIEW** ................................................................ 13

- Introduction ............................................................................................................ 13
- Assessment ............................................................................................................. 14
  - Reasons for assessment ....................................................................................... 16
  - Types of assessment ............................................................................................ 19
  - Stakeholders in assessment .................................................................................. 20
  - The assessment process ....................................................................................... 22
  - Assessment summary ........................................................................................... 34
- Electronic performance support system .................................................................. 34
  - Human performance improvement ...................................................................... 35
  - Knowledge management ...................................................................................... 39
  - Computer-assisted assessment ........................................................................... 40
  - Summary of review on EPSS ............................................................................... 42
- Design and development of an EPSS ..................................................................... 42
- Characteristics of the proposed EPSS ..................................................................... 46
- Summary ................................................................................................................ 47

**CHAPTER THREE THEORETICAL FRAMEWORK** .................................................. 49

- Introduction ............................................................................................................. 49
The human performance technology framework ...................................................... 51
The task assessment framework .............................................................................. 56
Performance-centered design framework ................................................................. 58
Summary ................................................................................................................ 60

CHAPTER FOUR METHODOLOGY ....................................................................... 61
Introduction ............................................................................................................ 61
Development of the research design ........................................................................ 63
Description of the research design .......................................................................... 66
  Exploration Phase ..............................................................................................................68
  Development Phase .........................................................................................................69
  Evaluation Phase ...............................................................................................................69
Target population and setting .................................................................................. 70
Data collection ........................................................................................................ 72
  Data collection during the Exploration Phase .................................................................73
  Data collection during the Development Phase ..............................................................74
  Data collection during the Evaluation Phase ...................................................................76
Data analysis ........................................................................................................... 77
Ethical considerations ............................................................................................. 79
Summary ................................................................................................................ 79

CHAPTER FIVE EXPLORATION PHASE DESCRIPTION AND FINDINGS ...... 81
Introduction ............................................................................................................ 81
Performance analysis .............................................................................................. 82
  Desired performance .........................................................................................................82
  Actual performance ..........................................................................................................87
  Performance gap ..............................................................................................................89
Cause analysis ........................................................................................................ 90
  Best practice interviews ....................................................................................................91
  Suggested intervention strategies .....................................................................................93
  Role of ICT ........................................................................................................................96
Conclusion .............................................................................................................. 97

CHAPTER SIX DEVELOPMENT PHASE – SEMESTER 1 DESCRIPTION AND FINDINGS ................................................................................................................ 99
Introduction ............................................................................................................ 99
Development Phase overview ................................................................................. 99
  Development platform .....................................................................................................100
  Team members ...............................................................................................................100
  Description of units ........................................................................................................102
Description of a typical development cycle ........................................................... 103
Semester 1 development cycles ............................................................................. 106
  Oral presentation development cycle (DC1) ............................................................... 107
    Description of assessment task ....................................................................................107
    Description of activities ...............................................................................................107
    Discussions of findings from DC1 ..............................................................................116
  Poster presentation development cycle (DC2) ..............................................................122
CHAPTER SEVEN DEVELOPMENT PHASE - SEMESTER II DESCRIPTION AND FINDINGS

Introduction ............................................................................................................ 145

Tutorial paper development cycle (DC4) ............................................................... 146
Description of assessment task .............................................................................. 146
Description of activities ......................................................................................... 147
Student response to Activities 1B and 5A ............................................................ 156
Discussion of findings from DC4 ........................................................................... 158

Event analysis map development cycle (DC5) ...................................................... 160
Description of assessment task .............................................................................. 160
Description of activities ......................................................................................... 161
Discussion of findings from DC5 ........................................................................... 166

Written exam development cycle (DC6) ............................................................... 167
Description of assessment task .............................................................................. 167
Description of activities ......................................................................................... 168
Discussion of findings from DC6 ........................................................................... 172

Conclusion ............................................................................................................ 173

CHAPTER EIGHT EPSS EVALUATION .................................................................... 175

Introduction .......................................................................................................... 175
Design themes ...................................................................................................... 177
Suggested additional design features .................................................................... 180
Usability themes ................................................................................................... 182
Efficiency construct of usability ........................................................................... 182
Effectiveness construct of usability ....................................................................... 184
Satisfaction construct of usability ......................................................................... 185
Usefulness construct of usability ........................................................................... 186
Implementation themes ......................................................................................... 187
Attitude of participants ........................................................................................... 187
Training .................................................................................................................... 188
ICT experience ....................................................................................................... 189
File management, version control and maintenance ............................................... 191
Distribution and collation of data ......................................................................... 192
Future directions .................................................................................................... 192
Summary ................................................................................................................ 193

CHAPTER NINE FINDINGS AND DISCUSSION .................................................. 195

Introduction .......................................................................................................... 195
LIST OF TABLES

Table 2.1 Conditions under which assessment supports student learning ................................. 19
Table 2.2 Purpose of assessment from the stakeholders’ perspective ........................................ 21
Table 2.3 Range of design parameters considered in the development of the EPSS .................. 46
Table 3.1 Phases of study related to HPT stages ...................................................................... 52
Table 4.1 Time frame and research questions as they relate to the phases of the study .......... 68
Table 4.2 Data sources, methods of collection and purpose by phase of the study .................... 73
Table 4.3 Links between research themes, phases and data collection methods ...................... 77
Table 5.1 Responsibilities of the unit coordinator according to ECU’s policy on assessment ... 84
Table 5.2 Responsibilities of the unit lecturers and tutors according to ECU’s policy on assessment ...................................................................................................................... 85
Table 5.3 Determination of marks and grades according to ECU’s policy on assessment ........ 86
Table 5.4 Performance gap range on key desired performance goals of the task assessment process ........................................................................................................................................ 90
Table 6.1 Unit team members .................................................................................................... 100
Table 6.2 Assessment tasks and development cycles by unit ...................................................... 103
Table 6.3 Tutorial oral presentation development cycle (DC1) .................................................. 108
Table 6.4 Example of tutors’ comments and coordinator’s revised grade descriptors (DC1) .. 111
Table 6.5 Poster presentation development cycle (DC2) .......................................................... 123
Table 6.6 Essay development cycle (DC3) ................................................................................. 132
Table 7.1 Tutorial paper development cycle (DC4) .................................................................. 148
Table 7.2 Tutorial groups grade descriptors with coordinator’s revised version (DC4) ............ 150
Table 7.3 Positive responses to the e-survey (DC4) .................................................................. 156
Table 7.4 Event analysis map development cycle (DC5) ........................................................... 162
Table 7.5 Exam assessment development cycle (DC6) ............................................................. 168
Table 8.1 Summary of the key design decisions and features by activity ................................. 179
Table 8.2 The development of the EPSS features .................................................................... 180
Table 10.1 Summary of student and tutor numbers to trial the EPSS post study ...................... 231
# LIST OF FIGURES

Figure 2.1 Complexity and importance of assessment...............................................................15
Figure 2.2 Relationship of assessment to learning. ...................................................................18
Figure 2.3 Key components of the assessment process...............................................................23
Figure 2.4 HPT model. Source: Van Tiem et al (2004). .................................................................37
Figure 2.5 Human Performance model. Source: Wile (1996). ......................................................38
Figure 2.6 User-Design Power Dynamics. Source: A. Carr (1997, p. 8). .....................................43
Figure 2.7 The curve of Pareto’s Law. Source: Nielsen and Mack (1994b). ................................44
Figure 2.8 Minimalist prototyping method. Source: Nielsen (1994a, p. 250). ..............................44
Figure 2.9 Cost-benefit trade-off curve. Source: Nielsen (1994a, p. 252). ...................................46
Figure 3.1 HPT Framework. Adapted from Van Tiem et al (2004). ..............................................52
Figure 3.2 Possible areas of intervention to improve performance. Adapted from Wile (1996). 54
Figure 3.3 The assessment process at the workplace level. .........................................................57
Figure 3.4 PCD Framework. Adapted from A. Carr (1997, p. 8). .................................................60
Figure 4.1 Location of research design. Source: Kumar (1999, p. 8). ...........................................63
Figure 4.2 User expectation of a product. Source: Vredenburg et al. (2002, p. 179). ...............65
Figure 4.3 Study design and phases.............................................................................................67
Figure 4.4 The task assessment development cycle. ....................................................................76
Figure 5.1 Example of a criterion and word levels.........................................................................88
Figure 6.1 Sequence of activities within the development cycle. .................................................105
Figure 6.2 Assignment 1 home page. ..........................................................................................110
Figure 6.3 Tutor online moderation marking of two criteria (DC1). .............................................112
Figure 6.4 Tutor marking view – A4 landscape (DC1). .................................................................113
Figure 6.5 Student feedback view – A4 portrait (DC1). ..............................................................114
Figure 6.6 Spreadsheet view (DC1). ............................................................................................115
Figure 6.7 Criterion where four radio buttons were used in a grade descriptor (DC1). ...............121
Figure 6.8 Previous poster marking key (DC2). ..........................................................................124
Figure 6.9 Poster presentation student feedback sheet (DC2). ...................................................126
Figure 6.10 Poster presentation tutor marking view (landscape layout) (DC2). ............................127
Figure 6.11 Pop-up menu for recording marks. ..........................................................................128
Figure 6.12 Essay assignment student feedback view (DC3). .....................................................134
Figure 6.13 Essay assignment tutor marking view (DC3). ...........................................................136
Figure 6.14 Grand total unit mark and grade (DC3). .................................................................137
Figure 6.15 Distribution of grades for one tutorial group (DC3). ...............................................139
Figure 6.16 Grade descriptor marks recorded as text. ...............................................................143
Figure 7.1 Previous assessment criteria (DC4). .........................................................................147
Figure 7.2 Tutorial paper student feedback view (DC4) ............................................................. 151
Figure 7.3 Peer group marking view (DC4). .............................................................................. 152
Figure 7.4 Tutor moderation marking view (DC4). ................................................................... 153
Figure 7.5 Spreadsheet comments view (DC4). ....................................................................... 154
Figure 7.6 Previous year’s marking key (DC5). ....................................................................... 161
Figure 7.7 Student feedback view (DC5). ............................................................................... 164
Figure 7.8 Peer group marking view (DC5). ............................................................................ 165
Figure 7.9 Summary view examples (DC5). ............................................................................ 166
Figure 7.10 Exam Q2 marking view (DC6). ............................................................................. 171
Figure 7.11 Unit spreadsheet view (DC6). ............................................................................... 172
Figure 9.1 Task assessment activities. ....................................................................................... 196
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQA</td>
<td>Assessment and Qualifications Alliance</td>
</tr>
<tr>
<td>AR</td>
<td>Action Research</td>
</tr>
<tr>
<td>AVA</td>
<td>Australian Volunteers Abroad</td>
</tr>
<tr>
<td>CAA</td>
<td>Computer-Assisted Assessment</td>
</tr>
<tr>
<td>CAL</td>
<td>Computer Aided Learning</td>
</tr>
<tr>
<td>CBT</td>
<td>Computer Based Assessment</td>
</tr>
<tr>
<td>CSHE</td>
<td>Centre for the Study of Higher Education</td>
</tr>
<tr>
<td>DC</td>
<td>Development Cycle</td>
</tr>
<tr>
<td>EA</td>
<td>E-Assessment</td>
</tr>
<tr>
<td>ECU</td>
<td>Edith Cowan University</td>
</tr>
<tr>
<td>EPS</td>
<td>Electronic Performance Support</td>
</tr>
<tr>
<td>EPSS</td>
<td>Electronic Performance Support System</td>
</tr>
<tr>
<td>HCI</td>
<td>Human Computer Interaction</td>
</tr>
<tr>
<td>HP</td>
<td>Human Performance</td>
</tr>
<tr>
<td>HPI</td>
<td>Human Performance Improvement</td>
</tr>
<tr>
<td>HPT</td>
<td>Human Performance Technology</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>ID</td>
<td>Instructional Design</td>
</tr>
<tr>
<td>ISO</td>
<td>International Standards Organization</td>
</tr>
<tr>
<td>ISPI</td>
<td>International Society for Performance Improvement</td>
</tr>
<tr>
<td>KM</td>
<td>Knowledge Management</td>
</tr>
<tr>
<td>LMS</td>
<td>Learning Management Systems</td>
</tr>
<tr>
<td>MCQ</td>
<td>Multiple Choice Questions</td>
</tr>
<tr>
<td>OLA</td>
<td>On-Line Assessment</td>
</tr>
<tr>
<td>PAR</td>
<td>Participatory Action Research</td>
</tr>
<tr>
<td>PBA</td>
<td>Problem Based Assessment</td>
</tr>
<tr>
<td>PBL</td>
<td>Problem Based Learning</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computers</td>
</tr>
<tr>
<td>PCD</td>
<td>Performance-Centered Design</td>
</tr>
<tr>
<td>PD</td>
<td>Participatory Design</td>
</tr>
<tr>
<td>QT</td>
<td>QuickTime</td>
</tr>
<tr>
<td>SMS</td>
<td>Student Management Systems</td>
</tr>
<tr>
<td>TBA</td>
<td>Technology Based Assessment</td>
</tr>
<tr>
<td>UCD</td>
<td>User Centered Design</td>
</tr>
<tr>
<td>VHS</td>
<td>Video Home System</td>
</tr>
</tbody>
</table>
Assessment is the process of collecting evidence and making judgements as to how well students have achieved the intended learning outcomes (University of Canberra, 2003).

Assessors are educators who are assigned the task of marking the assessment activity.

E-learning or Technology-Based Learning covers a wide set of applications and processes, such as Web-based learning, computer-based learning, virtual classroom, and digital collaboration. It includes the delivery of content via Internet, intranet/extranet (LAN/WAN), audio/video tape, satellite broadcast, interactive TV, and CD-ROM. (Cross, 2001; Hambrecht, 2000)

Electronic Performance Support System (EPSS) provides electronic task guidance and support to the user at the moment of need. EPSS can provide application help, reference information, guided instructions and/or tutorials, subject matter expert advice and hints on how to perform a task more efficiently. An EPSS can combine various technologies to present the desired information. The information can be in the form of text, graphical displays, sound, and video presentations (Gery, 1997).

Examination is defined as a test, quiz, essay paper etc set for students to complete in the official examination period at the end of each semester and which is invigilated (University of Canberra, 2003).

Intellectual Capital (IC) is a second order multi-dimensional construct. Its three sub-domains include:

i) Human Capital – the tacit knowledge embedded in the minds of the employees;

ii) Structural Capital – the organisational routines of the business, and

iii) Relational Capital – the knowledge embedded in the relationships established with the outside environment (Bontis, 1999).

Knowledge Management (KM) refers to capturing, organising, and storing knowledge and experiences of individual groups within an organisation and making it available to others in the organisation (Galbreath, 2000; Hambrecht, 2000).

Moderation is a process that is used to help assessors to arrive at a shared understanding of standards and expectations (HM Inspectors of Schools, 1999).
QuickTime is an application developed by Apple Computer in 1991, that is part of the system software architecture that seamlessly integrates audio, video, and animation across applications (Hansen, 1999, p. 256).
CHAPTER ONE
INTRODUCTION

Assessment is one of the core business activities of the tertiary education sector and tends to determine what is taught and learnt. It is a complex activity, which makes significant demands on the time, resources and emotions of learners and staff. The current research literature on assessment practice in higher education indicates a high level of disquiet and concern (Skidmore, 2003). This disquiet and concern is in part due to the extreme and rapid changes in higher education that have taken place since the early 1980s, and it implies the need to improve the assessment literacy of all stakeholders – learners, tutors/assessors, staff, the institution, public and business (Biggs, 1999; S. Brown & Glasner, 1999; Gibbs & Simpson, 2004).

University education in Australia has undergone dramatic and rapid changes in teaching and organisation since the early 1980s, within an increasingly complex educational and social environment. These changes have included: an increase in the size and cultural diversity of the student population, including a trend towards the internationalisation of the student population; the corporatisation and market orientation, along with a decrease in government funding; an increased expectation of the quality of education and institutional performance accountability and litigation; an increased access to and use of information and communication technology (ICT); a growing complexity and flexibility in the delivery of education; and the outsourcing of jobs with an increase in casual and contract employment (S. Brown, Thompson, & Armstrong, 1997; Choat, 2006; Cranton, 1997; DEST, 2002; Gibbs, Farrell, & Pollard, 1994; Nation & Evans, 2000; Taylor, 1999).

Prior to the early 1980s, units of study were year-long, often with only an exam as the assessment point and with one assessor; class sizes were small; the student population was more homogeneous; and lecturers tended to work from a craft knowledge approach to teaching and assessment (Speck, 2002). Since then, teaching and learning have become more complex, involving an increased use of ICT, large class sizes, multiple assessment points and assessors, and units of study that are one semester long. Additionally, the student population have become more diverse, multicultural and international, and university life is no longer the main focus for students (S. Brown et al., 1997; James, McInnis, & Devlin, 2002). These changes and others to university
education have affected teaching and learning programmes in the field of adult learning (androgogy), especially in the area of instructional design and methods (Knowles, Holton, & Swanson, 2005). For example, they have led to an emphasis on efficient and effective teaching and learning methods, such as problem-based learning and competency-oriented learning, and a global trend towards criterion- and standards-based curricula. These alternative, and arguably more authentic, teaching and learning methods use rich, meaningful and realistic assessment tasks that require the learner to demonstrate deep learning.

These assessment methods require a greater use of educators’ androgogical content knowledge (Lederman & Gess-Newsome, 1999) and professional judgement (Falk & Ort, 1998), and require new forms of marking, grading and reporting (O'Donovan, Price, & Rust, 2004; Rust, Price, & O'Donovan, 2003; Wiggins, 1998; Winter, 2003). These changes in focus for assessment represent a major shift from students’ factual recall or shallow learning to a demonstration of their knowledge and understanding, or deep learning. At the same time there is demand for greater levels of reliability (consistency), validity, accountability and transparency of the assessment process (Broadfoot, 1996; S. Brown & Wisker, 1996; Nation & Evans, 2000; Taylor, 1999). A discrepancy between actual practice and good assessment practice has in part contributed to the disquiet and concern found in the literature (Orrell, 2006).

Another significant educational assessment trend has been the expansion of the purpose of assessment from assessment of learning, that is, marking, grading and accreditation, to educative assessment, that involves assessment for and as learning. Biggs (1999), Black (2000), Brown (1999), Boud (1995a), Ramsden (1992), and Winter (2003), among others, have shown the significance of this new view of assessment for teaching and learning. This trend has also highlighted the shortcomings of the current assessment processes and practices involved in teaching and learning at all levels of education.

These trends and issues, and the difficulties of implementing these new forms of assessment, have been a catalyst for further research into assessment practice in higher education, and have led to many recent Australian and international conferences and reports on this topic. One example was the Centre for the Study of Higher Education’s (CSHE) publication Assessing Learning in Australia, commissioned by the Australian Universities Teaching Committee (James et al., 2002). This report focused on student assessment within Australian universities in 2002, and one of its findings was that
assessment practice was still often treated merely as the end-point of the teaching and learning process, that is, as assessment of learning. This concern on assessment practice also culminated in the first Australian conference solely devoted to assessment and evaluation held in 2002, and the First International Conference on Enhancing Teaching and Learning through Assessment held in Hong Kong in 2005 (Frankland, 2005).

The use and application of ICT by administration, staff and students has been one of the major changes that has occurred in the recent past in teaching and learning in Australian universities, and is continuing to expand rapidly. Although this change has markedly affected teaching and learning for both lecturers and students, assessment has been least affected. As Gipps (2005) states “in universities, the use of ICT in learning and teaching is much further advanced, while the use of ICT to support assessment is more patchy” (p. 172). Examples of advancement on the student side of the assessment process include written work required to be word-processed, the use of Microsoft PowerPoint in oral presentations, and the use of electronic reference material. Whereas on the staff side of the assessment process, the way assessments are marked, reported and managed, has been hardly affected by the growing access to ICT (Clyde & Delohery, 2005; Maier & Warren, 2000). In other words, the application of ICT to the staff side of assessment is still in its infancy, especially where professional judgement is involved.

In summary, many of these internal and external changes to university education are leading towards an ICT or digital educational environment (Reeves, 2002), that is, an integrated e-learning approach to education (Van Merriënboer, Jochems, & Koper, 2004), sometimes referred to as technology-enhanced and web-based learning. This evolving digital transformation of teaching and learning requires the integration and collaboration of different domains, disciplines and fields of study, such as instructional design (ID), human computer interaction (HCI), software design and development, and human performance (HP). Each of these disciplines has its own concepts, tools, methodologies and terminology, and although there are some similarities between them, the differences must be acknowledged and appreciated when researching across these fields.

**Aim**

The aim of the study was to investigate, develop and evaluate the application of electronic performance support (EPS) and strategies to enhance and improve the
performance of the task assessment process in university courses where professional judgement is involved in assessing the task. This study took a broad definition of task assessment that included all activities carried out by staff before, during and after marking the task set for assessment. These activities ranged from the marking key design to management, reporting and administration aspects, including quality control and assurance.

In essence, the study involved combining innovative assessment strategies with the movement of the marking sheet and associated processes off the desk and onto the desktop (computer screen), and investigating how this might augment, enhance and improve the performance of the task assessment process.

**Rationale**

A number of critical thresholds have come together to underpin the relevance and importance of this study. The growth in access to and improvements in ICT services has enabled this emergent area of digital assessment or e-assessment (JISC, 2007) research, but this growth is not sufficient justification for the investigation and implementation of digital assessment approaches. The study is justified when this growth in ICT is combined with the following areas of change and concern:

- Academic interest in using ICT to improve teaching and learning;
- Academic interest in using ICT to improve productivity;
- Increasing staff workload;
- Casualisation of staff;
- Increase in class sizes;
- Changing methods of assessment;
- Increased emphasis on appropriate assessment;
- Diversity of the student population; and
- Growing awareness that assessment is more than accreditation.

Access to ICT facilities is growing continuously, and includes access to desktop and laptop computers, printers, networks, wireless technology, the Internet and email. For example, at Edith Cowan University (ECU), the proportion of staff with an allocated laptop computer has grown from 11% in 2001, to 22% in 2004, and 52% in 2005. However, while ICT facilities and resources in higher education are becoming ubiquitous, staff use and productivity have not kept pace with this growth in access. This is no more so than in the area of assessment, where the application of ICT has been
minimal (Bottino, 2004; Van Merriënboer et al., 2004). This lag mirrors the situation that the corporate sector faced between 1978 and 1996, when after increasing their investment in technology by 600%, they found no similar increase in productivity (Stolovitch, 2000).

Nevertheless, the increased access to digital technology in universities is changing the traditional learning and teaching paradigm, and leading to an integrated e-learning approach (Van Merriënboer et al., 2004). This change began with analogue technology and has been under way for a long time, but is finally reaching a tipping point, where all educational transactions now pass at least through one digital media stage. Moreover, educators now have e-learning and e-assessment (McFarline, 2001; Ridgway & McCusker, 2004) as discipline fields of study, research and practice. As computers, that is, digital technology (e- is often used to indicate this) and Internet access become a significant part of the educational environment, computer-assisted assessment (CAA) (Bull & Sharp, 2000; Philip Race, Bull, & Brown, 1999) and computer aided learning (CAL) are also growing in significance.

Gipps (2005), found that assessment practice and process have tended to be unbalanced and seem to be lagging behind the application of research findings in teaching and learning practices. Brown and Knight (1994), in discussing the importance of assessment, found it was “still not the high-profile issue it should be” (p. 46). This is especially the case when one considers how learners and assessors use ICT to carry out the assessment process. While students are expected to use ICT in their assessment task, for example, to word-process their assignments, use slideshows for presentations, engage in discussion board interactions, engage in computer-based group-work, and use email, paradoxically, lecturers typically only use the most basic features of ICT applications available to them during the assessment process. For example, the researcher has observed that most staff at his University use a word processor only to create the mark-sheet templates on which they manually record student details, marks and comments. Similarly, when a spreadsheet application is used to produce a list of students and marks, many staff manually record marks and tally them. These observations support the findings of a wide-ranging survey of academic use of ICT carried out by Jankowska (2004), who found that only “24 percent used computers to evaluate students’ work” (p. 54).

While there is some literature about the use of generic application tools (e.g., word processors, spreadsheet and presentation applications) (Jankowska, 2004) and
learning management systems (LMS) (e.g., WebCT and Blackboard) by staff, less is known about the use of electronic performance support systems (EPSS) that focus on teaching and assessment. CAA is a sub-set of the much larger field of EPSS, while EPSS is a sub-set of the research field of human performance improvement (HPI). CAA has the potential to improve the performance of many aspects of assessment for all educational stakeholders. Currently, however, the focus of CAA has been on objective-based questions, with the automation of the whole assessment process by the computer, from the management to the marking of the questions. Thus, staff are left out of the assessment process by these current CAA methods of task assessment. The researcher believes there is a need for the development of appropriate CAA in the form of EPSS applications that do not just automate the task assessment methods, usually in the form of multi-choice questions, but involve both staff and students in the task assessment process. This study proposed to design, develop and evaluate an EPSS that would bring the staff back into the CAA task assessment process and provide them with electronic performance support. Thus, this research would contribute to the improvement of performance and practice of staff and learners in task assessment where professional judgement is involved.

The application of EPSS has grown over the last fifteen years in the business world with the inclusion of knowledge management (KM) (Massey, Montoya-Weiss, & O'Driscoll, 2005; McManus & Rossett, 2006). The literature in these two fields of KM and EPSS suggests to the researcher that educators’ professional knowledge and judgement could be captured and used in the assessment process. By incorporating KM into the design of the marking key, the valuable and useful tacit knowledge of experienced assessors (even when they no longer teach the unit) and that of moderators could be captured and made available to both learners and tutors (assessors) and to future coordinators of the unit. This knowledge, when accessed electronically via an EPSS, could be considered a cognitive and teaching aid to support the moderation, marking and management processes. These types of electronic performance aids could then further help in the training and professional development of learners and neophyte assessors.

A further catalyst for the study was the researcher’s knowledge and interest in the productive and educative application of ICT to teaching, learning and assessment at the unit and course level of study. The researcher’s personal teaching experience and observation of other lecturers and students using ICT over many years led him to
conclude that ICT has been and still is under-utilised in all areas of education. In particular this appears to be the case in the areas of marking key design, marking, moderation, feedback, reporting and management processes. The limited research evidence seems to indicate that assessment is currently very time-consuming, costly and stressful for both learners and staff when professional judgment is involved. While the clerical and administrative paperwork involved in assessment is important, time-consuming and complex, it often distracts and hinders good marking practice. These assessment practices and processes currently involve very little use of ICT (McFarline, 2001). When ICT is used in assessment, it seems to be neither integrated, nor linked across the processes or activities. This is especially so where professional judgement is involved and when more than one assessor is involved.

Freeman and Lewis (1998), in their discussion on “workload at the various stages of assessment” (p. 296), could find little information on either how long a student should spend doing the task nor on how long an assessor should take to mark it. Moore, Orey, and Hardy (2000) completed one of the few detailed task analysis of educators’ work activities at a high school. They found that during a normal teaching day, teachers spent 20.3 minutes of school time and 27.0 minutes at home assessing or marking student work. Thus, just in terms of time, assessing is an important component of educators’ activity, but when all the other aspects of assessment, as discussed earlier, are considered, the possibility of applying an EPSS to improve the performance of this important educational activity becomes compelling.

**Significance**

This study identified a significant performance gap in the task assessment process in tertiary education, even though at the institutional level a considerable amount of resources have been allocated to ICT infrastructure, access and support, and student management systems (SMS) over the last 20 years. These SMS include the availability of electronic unit material, the electronic communication with students, and the use of electronic curriculum and student management products such as learning management systems (LMS). However, there has been little regard for or research into how this increased availability of ICT for teaching, learning and assessment has actually benefited the stakeholders at the workplace level in terms of increased productivity or performance, or how this productivity might be measured or achieved. This is particularly the case in the assessment area, which appears to have been the most
neglected in terms of application of ICT (Gipps, 2005). This study investigated the application of strategies with a focus on electronic performance support systems (EPSSs) that would take advantage of this increased access to ICT at the workplace level to reduce the significant performance gap in the task assessment process, as had previously been achieved in the corporate sector.

Currently assessment activities in most Australian universities are mainly paper-based, or may pass through a digital phase but end up being printed for assessment purposes. The marking, recording and management systems employed in the assessment process in universities have been identified by the researcher as being areas where improved performance could be achieved through the application of an EPSS. Furthermore, the incorporation of KM into the EPSS could alleviate to some extent the growing shortage of experienced educators and the loss of their tacit knowledge of assessment criteria.

This study investigated a particular aspect of assessment, that of the task assessment at the workplace or course unit level, and involved the actual workflow practices of the stakeholders involved in high stakes assessment where professional judgement is required. The research literature on assessment practices highlights the need for more efficient and effective professional development, quality control and assurance procedures (Biggs, 1999; S. Brown & Glasner, 1999; Freeman & Lewis, 1998) and covers such issues as: validity, consistency of judgement (reliability), marking key design, transparency in marking, feedback, reporting, and management. Significantly, electronic performance support has the potential to address many of these performance issues. Combined with the critical threshold that has been reached in regards to access to ICT facilities in higher educational institutions, the feasibility of this type of cognitive tool to be used in the assessment process has now become an achievable goal.

The business sector has shown for over 15 years that where EPSS has been applied to workplace tasks, significant performance gains have been achieved (Gery, 1997; McManus & Rossett, 2006). Therefore, it seems reasonable to expect that similar performance gains could be achieved in education, specifically in universities and in the field of task assessment. These performance gains might be achievable in areas such as accountability, transparency, reliability, validity, moderation, marking, feedback, reporting and management. As Gipps (2005) stated, the “application of this technology could bring improvements in reliability and accuracy of marking, eradicate clerical
errors, speed up the marking process and, possibly, reduce the cost” (p. 172). The performance gains could also result in reduced stress and workloads for teaching staff involved in the assessment process.

The application of EPSSs could also enable micro- (sub-mark) and macro-analysis (between assignments) of individual student’s marks, resulting in improved feedback for student, tutor and coordinator, and allowing targeted support for specific students. Greater participation of both tutors and students in the task assessment process could also be achieved once the task assessment process had incorporated an EPSS. This participation by students could include the development and moderation of the marking key, self- and peer-group marking. Under a manual paper-based system, many of these innovative learning and assessment strategies have been difficult, time-consuming and often not economically feasible.

Scope of study

The scope of the study was the exploration of the design, development and use of an EPSS to augment, enhance and aid the performance of the task assessment process at the workplace level. The study was limited to university level task assessment where professional judgement was involved. The initial proposal was to focus on pre-marking moderation activities, but this was broadened to include the whole task assessment process. The scope of the study, however, was restricted to areas or activities within the task assessment process that the researcher and team members had control or influence over. The selection and description of the team members is discussed in detail within the Methodology Chapter (Chapter 4) in the section Target population and setting (p. 70).

The EPSS also needed to be designed and developed, and not just implemented and evaluated. This was necessary as no existing task assessment EPSS could be located that was sufficiently flexible and modifiable. In designing and developing the EPSS, a number of educational factors had to be addressed, such as the transparency, reliability, and validity of the marking key, marking activity, quality assurance and control, and management activities.
As outlined above, the approach to the assessment of learners is changing at all levels of education, whether the assessment is formative or summative in nature. There is a move away from objective assessment tasks that usually involve shallow learning and a move towards more authentic, educative, subjective and higher order thinking assessment tasks that involve deep learning. At the same time, the requirements and demands for authenticity, accountability, reliability, validity and transparency in the assessment process are increasing for all stakeholders. The actual practice and achievement of these changes, requirements and demands has been difficult and limited, as the following two quotes indicate:

Assessment sometimes appears to be, at one and the same time, enormously expensive, disliked by both students and teachers, and largely ineffective in supporting learning. (Gibbs & Simpson, 2004, p 11)

In spite of the central importance of assessment in the work of universities, and the hundreds of years over which universities have been carrying out assessments, the current literature displays remarkable disquiet. (Winter, 2003, p. 112)

The research question and subsidiary questions evolved throughout the research study, as is consistent with the nature of action research, and participatory product design and development. Initially the focus was on moderation, but this expanded to include the whole task assessment process from the development and quality assurance of the marking key, to the marking activities (feedback, reporting and management), and the knowledge and skill of the assessor. The researcher found as the study progressed that all these activities were integral to the complete task assessment process.

The research question that guided the study became:

To what extent does the application of an electronic performance support system (EPSS) enhance and support the performance of the task assessment process: the management, reporting, marking key development, marking, feedback and moderation processes, where professional judgement is required in the task assessment of student work in a university course of study?

The study addressed this research question from two different perspectives. Firstly, the EPSS was evaluated by considering a number of subsidiary questions that
focused on three themes: design, usability, and implementation. These subsidiary questions are presented here within these three themes.

**Design theme questions:**
- What are the key features of an EPSS designed for the task assessment process?
- What do users regard as the advantages of the EPSS over the manual paper-based methods typically used in the task assessment process?

**Usability theme questions:**
- What components or features of the EPSS do users find useful?
- What common pattern of usage of the EPSS was observed?

**Implementation theme questions:**
- What are the constraints or obstacles to the effective use of an EPSS and supporting strategies for this type of process?
- What effect do the EPSS and supporting strategies have on the marking key development, marking, moderation, reporting and management processes?

Secondly, to further elicit the findings, a number of assertions were developed based on the summative interviews of the individual team members, and observations made by the researcher.

**Thesis overview**

This chapter introduced some of the significant changes that have occurred in teaching, learning and assessment over the last 25 years, identifying assessment as significant and under-researched at the practitioners level within the learning environment of higher educational institutions. Also identified was the lack of integrated ICT applications in the form of EPSSs to support the performance of the task assessment practice where professional judgement is involved. The aim, rationale, significance and scope of the study were also discussed. The chapter concluded with the statement of the research propositions, including the research question and subsidiary theme questions.

The following chapter, Chapter 2 (*Literature Review*), presents a critical review of the existing literature around three diverse disciplines: assessment, human performance and software design, which helped frame and inform the research study. Chapter 3 (*Theoretical Framework*) discusses the theoretical and conceptual framework of the
study developed from the literature review, setting the study in a normal and authentic educational environment. In Chapter 4 (Methodology), the principles of participatory action research (PAR), user-centered design (UCD) and human performance technology (HPT) are investigated and applied to inform the development of the methodologies selected in the study. The selected methods of investigation placed the tutors and coordinator, that is, the team members, in control and at the centre of the iterative design process.

The incorporation of an EPSS into the task assessment process was explored, developed and evaluated over 18 months. The first six months explored and analysed the current and desired performance of the task assessment process within the work environment. In Chapter 5 (Exploration Phase: Description and Findings), the findings and results from this exploratory phase are discussed. Chapter 6 (Development Phase: Semester I Description and Findings) and Chapter 7 (Development Phase: Semester II Description and Findings) describe the EPPS development cycles or iterations over each semester, and discuss the findings as they relate to each cycle. Chapter 8 (EPSS Evaluation) discusses the summative evaluation of the study as it relates to the EPSS, while Chapter 9 (Findings and Discussion) discusses the findings as they relate to the overall task assessment process, presenting them as assertions that emerged from the summative evaluation. Finally, Chapter 10 (Conclusions) presents an overview of the findings, recommendations, as well as the implications of the study and suggestions for further research.
CHAPTER TWO
LITERATURE REVIEW

Introduction

This chapter provides a structured summary of the literature reviewed to develop a conceptual framework and research design. Three distinct and diverse fields from the literature were identified and reviewed to help conceptualise, inform, develop and position this study. Firstly, the field of assessment, with specific reference to the task assessment process. Secondly, the human performance improvement (HPI) field (Sanders & Ruggles, 2000), with specific reference to electronic performance support systems (EPSS), and finally the field of software product design and development, with specific reference to methodologies involved in the successful development of useful and productive software. The investigation of this last field of research was necessary as there were no appropriate EPSS tools available for use in the study and therefore the study included the design and development of an EPSS prototype. These three distinct fields of research: assessment, performance support, and software design and development are explored and discussed in this chapter as they relate to the study. The conceptual framework developed from this literature review is described in the next chapter, Chapter 3.

The first section of this chapter reviews the literature on assessment and highlights its complexity and importance in education. The implementation of assessment involves theory, research, policy, and practice, all of which relate to the quality of teaching and learning. As all these factors evolve, there is a need for assessment to adapt and evolve. These changes have been reflected in Australian universities, where the educational environment and, thus, the requirements of assessment, have changed significantly over the last twenty years (Hinett & Knight, 1996), as discussed in Chapter 1.

The second section reviews the literature on electronic performance support systems (EPSS) and its relationship to the HPI field, and describes the Human Performance Technology (HPT) model that is used to analyse performance gaps and implement strategies to improve performance. The difference between training and performance support is also discussed, with specific reference to EPS (Gery, 1997).
Additionally, the expansion of the EPSS field to include knowledge management (KM) is raised (Santo, 2005; Schwartz, Divitini, & Brasethvik, 2000; Winslow & Bramer, 1994), as well as the link between EPS and computer-aided assessment (CAA).

The third and last section reviews the literature on software product design and development, with specific reference to participatory design (PD) and user-centered design (UCD), and the application of the HPT model. The literature in this area is part of the rapidly evolving field of human computer interaction (HCI), a large field of study concerned with the joint performance of tasks by humans and computers. This section explores the most appropriate methodologies that could be used to achieve the aims of the study and thus informs the proposed research question.

The three fields of research above, although important, are still under-researched and in their infancy. These fields have changed and evolved significantly over the last twenty years in focus, policy, research and practice. This has meant that, at times, technical terms used within them have not been precisely defined, have changed over time, or have been used inconsistently in different fields of research. Furthermore, the practice in these fields has often not been in alignment with the policies and/or theories, while the exchanges of concepts and ideas within and between these fields of study have been limited. The researcher found this a significant barrier to understanding when reviewing the existing literature across and within these fields. However, time and space does not permit more than the flagging of these important issues and concerns.

Assessment

Assessment practice is a deeply complex phenomenon that defines educational goals and outcomes and shapes student learning. Assessment processes make profound demands on students and teachers alike in terms of time, resources and emotions. (Orrell, 2005, p. 17)

As the quote by Orrell highlights, assessment is a significant and complex phenomenon, and it is a major component of any learning environment for all stakeholders. Assessment covers more that just educational issues, but also issues of the time, resources and emotions of all stakeholders involved. These three aspects of time, resources and emotions are under-represented in the literature on assessment.

The complexity of assessment can be viewed and investigated from many perspectives; one such perspective is shown in Figure 2.1, which presents the researcher’s exploration of the concept. Freeman and Lewis (1998) highlighted that
“there is much imprecision in the way assessment language is used” (p. 2), while Black and Wiliam (1998) found this same imprecision in the area of assessment and classroom learning in their comprehensive meta-analyses of the literature covering 681 publications. This imprecise language usage covers many of the assessment terms listed in Figure 2.1. Many of these terms are often inconsistently used or defined.

Figure 2.1 Complexity and importance of assessment.

The term assessment itself is not tightly defined and does not have one widely accepted meaning. For example, to perform their comprehensive review, Black and Wiliam (1998) had to physically turn the pages of seventy-six of the most likely journals in the field of assessment, as the inconsistency in the use of terms meant that the use of key-words was an inadequate tool in the literature search. Black (2000) concluded from this review process that “it seems most researchers are not studying much of the literature that could inform their work” (p. 408). Similar conclusions have been reached by other researchers, including this researcher when carrying out his own review of the assessment literature. For example, Miller, Cox and Imrie (1998) commented that “some educationists do not distinguish between assessment and evaluation” (p.3). The authors defined assessment concisely as any “means by which students’ progress and achievement are measured, recorded and communicated to students and relevant university authorities” (p. 4). In contrast, Dietel et al. (n.d.), defined assessment more expansively as:
Any method used to better understand the current knowledge that a student possesses. ... This implies that assessment can be as simple as a teacher's subjective judgement based on a single observation of student performance, or as complex as a five-hour standardized test. ... Assessment may affect decisions about grades, advancement, placement, instructional needs, and curriculum.

These definitions reflect the fact that there are many issues and questions still unresolved involving theory, policy and practice in assessment. All these unresolved issues and inconsistencies become apparent when one begins to research and reflect on assessment, whether this involves research on the different types of assessment – high or low stakes, summative or formative, norm-referenced or criterion-referenced, objective or subjective – or on the main stakeholders in assessment (e.g., student, educator, university, etc). All these associated assessment terms, including the term assessment itself, tend not to be tightly defined in the literature, and the relationships between the terms are rarely mutually exclusive. For example, at university, assessment tasks are often seen as having the dual role of being both summative and formative. As this study was limited to the task assessment process, with the focus on the performance improvement of the tutors and coordinator involved in this process, only the literature relevant to this area will now be discussed.

The majority of research on the improvement and enhancement of the assessment process (Freeman & Lewis, 1998; A. Miller et al., 1998; Wiggins, 1998) at all educational levels (Biggs, 1999; Black & Wiliam, 1998) has tended to focus on the improvement of the assessment task and feedback. That is, what the student has been asked to produce or perform for assessment and the feedback from the assessor. However, when the advantages of criterion-based marking keys (or the type and method of feedback to improve the marking process) are referred to in the literature, it tends to be in broad general terms. Research on assessment has been mainly in the areas of policy and theory about the assessment task (the student side of assessment), with less research carried out in the area of practice and especially performance of the task assessment (the staff side of assessment). The following section covers the reasons for, and types of assessment, and the perceptions of stakeholders involved in assessment, and concludes with an analysis of the assessment process at the workplace.

**Reasons for assessment**

*The main purpose of assessment is to discover if students have achieved the learning outcomes of the course studied. The term assessment is derived from the Latin phrase ad sedere: to sit down beside. Primarily then assessment should provide guidance and feedback to the learner. (Bone, 1999, p. 3)*
An effective assessment system needs to cover a range of reasons or functions for engaging in the process of assessment. Generally, these reasons are: to support learning, and provide feedback to learners, parents and other educators; to identify the next steps in learning; and to provide information as a basis for selection and certification (Freeman & Lewis, 1998). According to Brown (1999, p. 47) assessment has six main functions: 1) capturing student time and attention; 2) generating appropriate student learning activities; 3) providing timely feedback which students pay attention to; 4) helping students to internalise the discipline’s standards and notions of quality; 5) marking, generating marks or grades which distinguish between students or which enable pass/fail decision to be made; and 6) quality assurance, that is, providing evidence for others outside the course (such as external examiners) to enable them to judge the appropriateness of standards on the course.

There has been a growing awareness and acknowledgement among educators, policymakers, and others of the influence that assessment has on curriculum (Biggs, 1999; Freeman & Lewis, 1998). Bone (1999) described assessment as “one of the most effective ways of changing how and what students learn” (p. 4). This takes places where the backwash effect of assessment is positive, that is, when assessment is aligned to the curriculum (Biggs, 1999). Educators are turning to alternative assessment tasks or methods as a tool to achieve educational reform (Ashcroft & Palacio, 1996; Boud, 1995, 1998; S. Brown & Knight, 1994), as they realise that changes to the assessment process are needed to reform curricula and instruction. However, assessment is still “under-discussed and, in most disciplines, an under-researched aspect of higher education” (Fry, Ketteridge, & Marshall, 1999, p. 58). The importance of assessment and the lack of professional discussion in universities highlighted by Fry et al (1999) led the University of Queensland to convene a conference for staff of local universities in 1998, under the theme of Effective Assessment at University (University of Queensland, 1998). The continuing recognition of the importance of assessment was noted in 2002 in the first Australian conference on Evaluation and Assessment, and in 2005 in the first international conference solely devoted to assessment and evaluation held in Hong Kong: Enhancing Teaching and Learning through Assessment (Frankland, 2005).

**Assessment and learning**

The relationship between assessment and learning is complex, and is sometimes viewed too narrowly as assessment of learning. However, this definition of assessment
as just *marking* or *grading* is changing to include assessment *for* and even *as* learning. This relationship between assessment and learning is illustrated in Figure 2.2. A narrow definition of assessment however has been one of the reasons for the growing literature highlighting the failure of the assessment process to achieve its full educative potential (Winter, 2003). If assessment does drive student learning, as the literature suggest (Biggs, 1999; Ramsden, 1992), then the system is failing if only assessment *of* learning is taking place.

![Figure 2.2 Relationship of assessment to learning.](image)

There is a reasonable volume of literature on how to achieve assessment *for*, and *as* learning, through constructive alignment (Biggs, 1999; Elwood & Klenowski, 2002; Skidmore, 2003) and what to set as assessment tasks (Wiggins, 1998; Winter, 2003) to achieve this. However, the literature on how to mark these assessment tasks is scarce, and often limited and/or very generic and global in nature. Take for example the latest offerings from Gibbs and Simpson (2003) of conditions under which assessment supports learning (see Table 2.1). These conditions provide some direction, however, they do not directly help the practicing teacher on how to apply them to their assessment practice.
Table 2.1
*Conditions under which assessment supports student learning*

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity and distribution of student effort</td>
<td>1. Assessed tasks capture sufficient study time and effort</td>
</tr>
<tr>
<td></td>
<td>2. These tasks distribute student effort evenly across topics and weeks</td>
</tr>
<tr>
<td>Quality and level of student effort</td>
<td>3. These tasks engage students in productive learning activity</td>
</tr>
<tr>
<td></td>
<td>4. Assessment communicates clear and high expectations to students</td>
</tr>
<tr>
<td>Quantity and timing of feedback</td>
<td>5. Sufficient feedback is provided, both often enough and in enough detail</td>
</tr>
<tr>
<td></td>
<td>6. The feedback is provided quickly enough to be useful to students</td>
</tr>
<tr>
<td>Quality of feedback</td>
<td>7. Feedback focuses on learning rather than on marks or students themselves</td>
</tr>
<tr>
<td></td>
<td>8. Feedback is linked to the purpose of the assignment and to criteria</td>
</tr>
<tr>
<td></td>
<td>9. Feedback is understandable to students, given their sophistication</td>
</tr>
<tr>
<td>Student response to feedback</td>
<td>10. Feedback is received by students and attended to</td>
</tr>
<tr>
<td></td>
<td>11. Feedback is acted upon by students to improve their work or their learning</td>
</tr>
</tbody>
</table>

**Types of assessment**

The types of assessment are just as complex and diverse as the purposes of assessment discussed in the previous section. Assessment has traditionally been divided into formative and summative, depending on how the assessment results are used (Bloom, Madaus, & Hastings, 1971; Brady & Kennedy, 2005; Weeden, Broadfoot, & Winter, 2002), and into norm-referenced (i.e. based upon discriminating between students) and criterion-referenced (i.e. based on judging whether students have met established standards) (Boud, 1998). While the assessment method or task might be the same, the reasons for and type of assessment may be different. For example, an essay could be either high or low stakes, assessed either formatively or summatively, and either norm-referenced or criterion-referenced.

Summative assessment is usually high stakes, is concerned with the ranking of learners, and is often used for passing course requirements. Formative assessment has been described as part of the natural learning process and is essentially diagnostic in nature (S. Brown & Knight, 1994). Boud (1998), in his paper titled *Assessment and learning – unlearning bad habits of assessment*, highlighted the move from norm-referenced to criterion-referenced assessment in tertiary education, noting that:

*Norm-referenced assessment is now prohibited by university policy at the University of Queensland and at an increasing number of other Australian universities. (Boud, 1998)*
Internationally, there is a similar trend towards a student-centred learning (SCL), and criterion-referenced, authentic and outcomes-based assessment (Black & Wiliam, 1998; Boud, 1998; Brady & Kennedy, 2005; S. Brown & Glasner, 1999). The term SCL has been described by Cannon and Newble (2000) as:

Ways of thinking and learning that emphasize student responsibility and activity in learning rather than what the teachers are doing. Essentially SCL has student responsibility and activity at its heart, in contrast to a strong emphasis on teacher control and coverage of academic content in much conventional, didactic teaching. (p. 16)

This trend has precipitated a movement away from objective-based assessment or assessment of explicit knowledge, to alternative assessment methods based on subjective or tacit knowledge assessment (O’Donovan et al., 2004). These methods of assessment include a variety of types of tasks such as open-ended questions, exhibits, demonstrations, hands-on experiments, writing in many disciplines, and portfolios of student work assembled over time. All these methods of assessment have one thing in common: they all require the educator to apply their professional or tacit knowledge judgement (O’Donovan et al., 2004) to determine whether the desired learning outcomes have been demonstrated.

Stakeholders in assessment

The reasons for and methods of assessment can both be viewed from the stakeholders’ perspective. Table 2.2, compiled from a number of sources (University of Canberra, 2003; University of Queensland, 1998), shows the different purposes of assessment from the perspective of all the stakeholders that have been identified in the literature on assessment in tertiary education: students, academic staff, the institution (university) and the broad community.

Research, although limited, has shown that workloads and stress amongst university staff (Hinett & Knight, 1996; Houston, Meyer, & Paewai, 2006) and students (Hughes, 2005) have increased due in part to changes in academic life. These changes have included the expanded purpose of university education to include a general educational experience, knowledge creation, vocational and employment preparation. The increase in accountability and litigation has increased staff stress and workloads; while the research on student stress has highlighted concerns about grades, relationship problems, loneliness, and money as major concerns for the majority of students. In two similar studies carried out 10 years apart (Furr, Westefeld, McConnell, & Jenkins, 2001;
Westefeld & Furr, 1987), the authors found that the most frequently reported concerns among students prior to a depressive episode were assessment grades or results. Adding to those findings, research has shown that the manner in which results are reported to learners can arouse negative emotions within them (P. Race, 1995).

Table 2.2
Purpose of assessment from the stakeholders’ perspective

<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Purpose of assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>to engage them in learning and to provide:</td>
</tr>
<tr>
<td></td>
<td>feedback on how effectively they are learning;</td>
</tr>
<tr>
<td></td>
<td>evidence that they have reached the required standard; and</td>
</tr>
<tr>
<td></td>
<td>evidence to show other people of their learning achievements.</td>
</tr>
<tr>
<td>Academic staff</td>
<td>to provide:</td>
</tr>
<tr>
<td></td>
<td>feedback on how effectively they are teaching;</td>
</tr>
<tr>
<td></td>
<td>evidence that their students have reached a particular standard;</td>
</tr>
<tr>
<td></td>
<td>evidence on what students know before commencing the subject; and</td>
</tr>
<tr>
<td></td>
<td>evidence to others that they are effective teachers.</td>
</tr>
<tr>
<td>The institution</td>
<td>to provide:</td>
</tr>
<tr>
<td></td>
<td>evidence that students have achieved learning outcomes;</td>
</tr>
<tr>
<td></td>
<td>evidence on how effective the teaching is;</td>
</tr>
<tr>
<td></td>
<td>evidence to show others that students have achieved what the institution claims they have; and</td>
</tr>
<tr>
<td></td>
<td>evidence that students have reached a particular standard.</td>
</tr>
<tr>
<td>The community</td>
<td>to provide evidence:</td>
</tr>
<tr>
<td></td>
<td>on what students have learned and the standards they have reached;</td>
</tr>
<tr>
<td></td>
<td>that graduates are employable; and</td>
</tr>
<tr>
<td></td>
<td>that institutions and their teaching programs are effective.</td>
</tr>
</tbody>
</table>

As academic staff adapt to these new curricula and assessment processes in countries such as the UK and Australia, there is growing evidence that there are three basic approaches being taken by educators towards assessment (Gipps, 1994b). Black (1993) defines three types of educators: *intuitives*, *evidence gatherers*, and *systematic planners* (assessors or integrators). In the *intuitive* approach to assessment, the educators rely on their memory, tradition and informal knowledge that are based on their tried and tested ideology that is centred on the learner’s needs. Black claims the *evidence gatherers* get lost in the gathering of evidence and see the collection of evidence as the end of the process. He claims that these new curriculum and assessment paradigms require *systematic planners*; these educators both systematically assess and integrate those assessments into the next set of learning activities to achieve the required evidence of the learning outcomes and, thus, are using assessment for learning.

Although describing educators’ practice in assessment, Black’s definitions also link in with educators’ (implicit) views of learning, and with their different attitudes and approaches to criterion-referenced assessment and formative assessment. Black’s
(1993) review found that only the systematic planners were correctly using formative assessment and, thus, correct professional judgement strategies. Thus, any marking and moderation process aiming to develop and improve the performance of educators’ knowledge and judgement will need to consider the current practices used by educators to carry out assessment (Samuelowicz & Bain, 2002).

The assessment process

The assessment process for the purpose of this study can be viewed as consisting of three components: 1) the assessment task (what the learner does); 2) the task assessment (what the assessor does); and 3) management and administration (what the stakeholders do with and how they receive the feedback and results).

As this study was concerned with the performance improvement and enhancing of the task assessment component through the application of an EPSS, this component was further broken down into the following three elements: a) the marking key (instrument of measurement); b) the skills and knowledge of the assessor; and c) the marking activities.

Figure 2.3 shows the researcher’s diagrammatic representation of the different components and elements of the assessment process and the relationship between them, based on the literature (G. Brown, Bull, & Pendledge, 1997; S. Brown & Glasner, 1999; S. Brown & Knight, 1994). The literature on each of these three elements of the task assessment is discussed below as they relate to the task assessment process. This section concludes with a discussion on the management and administration process (all these areas are highlighted in grey in Figure 2.3).

Some of the aspects of the task assessment process have been neglected in research at all levels of education, as Freeman and Lewis (1998) found in their comprehensive book Planning and Implementing Assessment, noting that “surprisingly little research has been carried out on workload” (p. 295), and this paucity of research applied to both students and staff. Meanwhile, in their research, Brown, et al. (1997) could not find any reliable figures on the time it took to do or mark assignments. This is a significant gap in research when one considers the importance of and the resources involved in assessment.
Figure 2.3 Key components of the assessment process.

The methods of assessment are numerous and a great deal of literature has been produced about them (G. Brown et al., 1997; Freeman & Lewis, 1998). The product or assessment task a student produces for assessment is dictated by the method of assessment. However, little literature has been published on how to actually carry out the marking and grading process, that is the task assessment. Often the literature seems to assume that the reader knows how to assess or mark the assessment task being described, discussed or evaluated. However, as the variety of methods of assessment discussed in the literature has been growing, for example group work, oral presentations or participation in discussion boards, discussion on how these new forms or methods of assessment are to be assessed or marked has been limited and poorly described.

The marking key

The marking key (assessment instrument) is the framework used to assess the assessment task. The assessment task is based on the assessment method used to demonstrate what has or has not been learnt. Marking key, feedback sheet, marking
scheme or check list are just a few of the terms used to describe the assessment instrument. The literature abounds with these terms but few are ever well described, justified or defined. Nor are the assessment methods, technique and implementation procedures well described in the literature. This lack of research into the quality and implementation of the marking key seems to be one of the missing links in the search for good practice in teaching, learning and assessment. Kuisma (1999) carried out one of the few comparative studies between a criterion-based marking key (though this was not an instructional rubric) and an intuitive one. He found that the criterion-based marking key achieved a greater spread of marks with improved objectivity, but was more time-consuming to carry out.

To achieve good practice in assessment, the marking key needs to be well thought out, clear, transparent and be aligned with the task and the learning outcomes (Biggs, 1999). Criterion- or rubric-based marking keys, when implemented correctly, offer the potential to achieve this good practice in assessment. However, as Rust et al. (2003) found in their two-year research project, tacit knowledge cannot be wholly articulated in text for either the assessors or the learners, confirming other similar research findings (O'Donovan et al., 2004; Saunders & Davis, 1998). These researchers found that strategies needed to be developed which went beyond clear and explicit criteria, if consistency, reliability and transparency were to be achieved between the stakeholders. The suggestions and strategies they developed included the following: the criteria should be jointly developed by the assessors and, where possible, learners; the criteria should be moderated and debated by the assessors and, where possible, learners, each time they are used; and procedures covering the task assessment process should be clear and articulated.

The word rubric is a derivative of the Latin word ruber meaning red. In literary history, rubrics are margin notes in texts giving description, or common examples for, or about, the passage (Wiggins, 1998). The current research literature on marking keys promotes the use of criterion- or rubric-based marking keys to enhance transparency, reliability and, when the task is aligned with the learning outcomes, also validity (Andrade, 2005; Coffin, 2002; Jackson & Larkin, 2002; McCollister, 2002; Montgomery, 2002; Rust et al., 2003; Tierney & Marielle, 2004). In current usage, a rubric is a guide listing criteria used for rating performance (Wenzlaff, Fager, & Coleman, 1999; Wiggins, 1998). However, some have argued that the use of the term
rubric needs to be more rigorously defined to allow for a more informed discussion (Wenzlaff et al., 1999).

Thus to distinguish between other forms of marking keys and what could be called a detailed rubric or set of criteria, the researcher applied the term rubric only to those marking keys where the levels or grades of achievement or performance were described. These achievement or performance descriptions are often called grade descriptors or descriptions (Greatorex, Johnson, & Frame, 2001). Thus marks or word grades indicating levels or grades without detailed descriptors would not be called a marking key rubric as they do not provide adequate transparency for either the assessor or learner. One reason that descriptors are not used is that to achieve this level of detail is not a trivial task and requires much thought and experience (Tierney & Marielle, 2004). However, the collaborative involvement of the assessors and learners, as Rust et al. (2003), O’Donovan et al. (2004) and Saunders et al. (1998) suggest, can greatly reduce this difficulty and improve transparency.

Assessment rubrics can take many forms and levels of complexity. However, a general description would be that they tend to use criteria that measure performance, behaviour or quality of output. These criteria contain a range of indicators, usually three to five, that are described in detail, showing the different levels of achievement that need to be reached to obtain a specific grade. For the purpose of this study an instructional rubric is the one that captures the most information about the assessment task and the learning outcomes using text (Andrade, 2005; Andrade & Boulay, 2003; Tierney & Marielle, 2004).

**Skills and knowledge of assessor**

As reforms based on standards sweep the country and educators grapple with ways to help an increasingly diverse student population realise its academic and social potential, the need for knowledgeable and highly skilled teachers becomes even more important. (Falk & Ort, 1998, p. 59)

The assessor, that is, the marker, is the person responsible for carrying out the task assessment. In the past in Australian universities, the assessor’s role was more stable, homogeneous, with smaller class sizes and with classes run over a year (Ecclestone, 2001). Programmes in the past were less diverse with less modularisation of units. This provided more time for understanding, learning and relationships between lecturers, who were also usually the assessors, and learners to develop. Currently, with increased modularisation and increased class sizes, the tutor and not the lecturer is usually the
In response to these changes, in particular the need to expand the number of assessors, Brown et al. (1999) suggest a number of other agents that could be considered as appropriate assessors: self, peer, employers and clients. However, regardless of who the assessor is, appropriate professional judgement is still critical in the assessment process.

In the past, and even currently in universities, assessment involving professional judgements has often been impressionistic (global), based on a connoisseur model of assessment (Webster, Pepper, & Jenkins, 2000) or “an elite ‘guild’ of professional assessors, whose professional judgement was mysterious in nature, and inaccessible to the layman” (O’Donovan et al., 2004 p. 326). These professional judgements were thus often made without using a detailed marking key and were usually norm-referenced. Feedback was not consistent and often consisted of ticks and hand-written comments that were difficult to read. Boud (1995), commenting on these problems of professional judgement, noted:

*It is assumed that there is always someone able to make a valid judgement of the matters under consideration. Disturbingly, the growing body of research on professional judgement casts doubt on the confidence with which we can hold this view. (p. 210)*

Based on the review of the literature by Plous on the psychology of judgement and decision-making, Boud (1995, p. 210) lists the following issues that are pertinent to task assessment that require professional judgement:

- Perception is selective;
- Commitment influences judgement;
- The wording of questions profoundly influences answers;
- Memory and hindsight bias judgement;
- Judgement depends on context; and
- Familiarity can offset context.

Although all these issues are important in forming an assessment judgement, the last three are the most relevant to this study. All assessors are liable to memory lapses and thus it is important to keep contemporaneous and accurate records and notes of assessors’ judgements. The context also affects judgement by the contrast, recency and halo effects. While familiarity can offset the effects of context, as familiarity decreases, assessors’ judgements become more easily influenced by context. Thus, the need for professional development, transparency, marking keys, moderation and other tools and
strategies to aid the performance of assessors in reducing the effects of these issues on
the assessment process can be seen (Elwood & Klenowski, 2002; Hinett & Knight,
1996; Rust et al., 2003; Saunders & Davis, 1998).

Marking activity

... the business of marking student scripts still remains as the most significant
quality event in the lives of the students and the academics. (Fleming, 1999, p. 83)

The marking activity involves different tasks depending on the type of assessment
being carried out. These tasks can include moderation of marks, marking of student
work and reporting of results. In high stakes assessment, marking is usually associated
with a mark or grade given to the learner, feedback often being given, and reporting of
these to authorities for the purpose of selection or certification and accountability
(Athanasou & Lamprianou, 2002). Marking and the tasks associated with it are not
often well defined in the research literature. For the purpose of this study, marking is
defined as the activity undertaken by the assessors, be they learners or instructors, when
they use a marking key to evaluate an assessment task. As this study was interested in
marking high stakes assessment that included a major degree of professional judgement
(i.e. a subjective component), the reliability, consistency and quality of the judgement
became crucial (Wiggins, 1993).

Bridges et al. (1999) found that depending on the degree of judgement or
subjectivity involved in the marking process, when a numeric marking system was used
at university they were able to categorise the marks into three types of distribution (A,
B and C). Type A distributions were associated with high levels of subjective
judgement (e.g., English and History) and were characterised by steep-sided negatively
skewed distributions with a narrow spread. That is, the range of marks was less in
qualitative or subjective type subjects. The second type, type B distribution, is slightly
broader spread and includes subjects like Biology, Business Studies, Language Studies
and Law. While in type C distribution (e.g., Computer Science and Mathematics), the
spread of marks was much wider. These are subjects where professional judgement
plays a lesser role in the marking process.

This distribution effect was attributed by Bridges et al. (1999) to the fact that once
professional judgement is required in the assessment process and a percentage marking
system is used, assessors find the awarding of either very high or low marks difficult to
justify. In other words, is there a difference between a mark of 63%, 65% and 67%? can
a marker justify this level of discrimination when using a non-descriptive marking key? The literature (see above section on marking key) suggests that the use of a detailed instructional rubric allows for a greater spread of marks over a greater range of grades by improving the transparency and the ability to justify the marks.

Another important aspect of the marking activity, apart from the use of a marking key, is feedback and reporting. Many definitions and terms have been used to cover feedback and reporting, such as *feedout*, *feedback* or *feedforward*. The term(s) used depends on the purpose of the assessment and are not usually mutually exclusive. Assessment has a feedout function (Knight, 2002a) when used in a certification process, and these types of assessment are called high stakes or summative. Meanwhile, the purpose of feedback and feedforward is to help teachers and learners, and guide further learning. There is often tension within the assessment process between the needs for feedout and feedback. Knight (2002a) even argues that where “feedout is the goal, disclosure is displaced by deception” (p. 277).

Carless (2006) is one of the few researchers to have investigated the perceptions of students and staff on feedback. He found that the differing perceptions were:

- Tutors believed that they were providing more detailed feedback than their students perceived;
- Tutors perceived their feedback to be more useful than perceived by their students;
- The perception of some tutors that students were only interested in their grades was brought into question. The balance of students focus on grades and/or future improvement remained unclear; and
- Tutors believed that their marking was fair, whilst students had mixed feelings about the fairness of the grading.

In addition, Carless (2006) found that students and staff held the following similar perceptions:

- Both tutors and students seemed to agree that students found it difficult to decode criteria; and
- Both tutors and students seemed cognizant of the emotional aspect of assessment.

All these perceptions highlight the need for more research in this area, given the importance of assessment in teaching and learning. The assessment process needs to be demystified by improving its transparency, and by educating students and staff about, and in, the process. This was demonstrated in research by Rust et al. (2003) reported in
the journal article *Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes*.

Finally, the marking activity (including marking, feedback and reporting) brings the spotlight on the emotional aspect of assessment for both learners and assessors, partly due to the authority relationship between them. Not only does marking, feedback and reporting need to be understandable, transparent, timely and acted upon by students (Gibbs & Simpson, 2004), but the psychological aspect of giving and receiving feedback is extremely important to both student learning and to assessors (Yorke, 2003).

**Management of assessment process**

Management and administration of the task assessment process involves all stakeholders and need to be made clear to learners and tutors. The following generic features need to be both clear and transparent: policy on deadlines and extensions; handing and method of confirmation; clear task requirements (word length, word-process, academic writing, plagiarism etc); turnaround time; and tracking of marks, students and quality of feedback (G. Brown et al., 1997; Freeman & Lewis, 1998). The literature on curriculum and assessment, as Yorke (1998) found, “has surprisingly little to say about the management of assessment” (p. 101), yet, the effective management of assessment is more important than ever with the changing nature of tertiary education, and the significant role it plays in the student experience and outcome of the assessment process. Institutions need to have in place effective assessment management plans. These plans need to be communicated to the stakeholders and need to cover student grievances, complaints and other types of litigation, as the student population grows and diversifies, and students perceive themselves as paying customers. Failure to do so could lead to unnecessary expense and litigation (Knight, 2002b; Yorke, 1998).

The issue of systematic management has implications from the institutional level down to the individual assessor, and covers the areas of accountability and transparency, coherence, equity, progress management, effectiveness and efficiency, and synergy within the institution (S. Brown & Knight, 1994; Knight, 2002b; Yorke, 1998). Graduate attributes and assessment policies and procedures are being devised and propagated in tertiary educational institutions, but without systematic management of assessment, implementing such policies will be problematic.
This study was interested in the fine grain issues of assessment management of individual task assessments and the combining of those individual assessments to produce a unit (course) mark. This involved the management and coordination of the task assessment activities and aspects discussed above, and the quality control and assurance aspects that are now discussed.

**Quality control and assurance**

Quality control and assurance of the assessment process at all levels needs to be undertaken for the process to be *valid*, *reliable* and *fair*. These terms, as most of the terms associated with assessment, tend to have a range of meanings. They have also tended to be defined and explained in terms of the assessment task or performance. The more traditional meaning of *validity* is “the extent to which a test measures what it was designed to measure” (Gipps, 1994a, p. 58). The use of the term *validity* has now expanded to include construct, content, consequential, concurrent and predictive validity, depending on the evidence collected and the purpose of the assessment (Brady & Kennedy, 2005). Furthermore, the discussion of validity has tended to focus on the assessment task and not on the task assessment process, which is the focus of this study.

*Reliability* of assessment refers to consistency in measurement of results between different assessors and test situations (Weeden et al., 2002). To achieve reliability two main approaches have been taken. One approach has involved the assessment task side, where test-retest, parallel forms and split-half procedures have been used (Athanasou & Lamprianou, 2002). A second approach has involved the task assessment side, where consistency in marking has been investigated. This has included multiple markings of one assignment by different assessors (inter-rater reliability) or the same assessor at different times (intra-rater reliability). This study was interested in improving both inter- and intra-rater reliability.

The term *moderation* is often used to describe the process used to achieve reliability in relation to grading and marking where multiple assessors or assessment tasks are involved. Currently the main moderation methods used in all levels of education are: professional discussion between assessors, the use of exemplars, and the use of clear assessment criteria (Brady & Kennedy, 2005). In tertiary education the tradition of moderation has been haphazardly applied, to the extent that often essay marking has been “traditionally marked by academic instinct” (Fry et al., 1999, p. 63). However, assessment not only needs to be reliable (accurate and consistent) and fair, it
also needs to reflect programme content (curriculum) and be valid (appropriate) (Freeman & Lewis, 1998). The following reasons are typically given for why assessment needs to be accurate and reliable (Fry et al., 1999; Preston & Shackelford, 1999): it is useful and fair to students; for internal and external quality assurance purposes; and to defend the increasingly likely legal challenge from disaffected students.

There is a growing need for justification and transparency in the allocation of grades and levels to satisfy learners, and the university and other official bodies. This places a heavy demand on educators and the assessment process to achieve reliability and validity (Freeman & Lewis, 1998). As Miller (1998) notes, “much needs to be done to develop assessment expertise” (p. 263), highlighting the need for “more formal training of staff in assessment techniques” (p. 263). Boud (1998), who has been publishing in the field of assessment for over 10 years, paints a bleak picture:

> My view is that a large part of what occurs in assessment in higher education is based on bad habits copied from the past and a lack of critical thinking about what we do. We are locked into patterns of assessment, which cannot be justified on any educational grounds whatsoever. Assessment is a vital part of teaching and learning. We should not undermine its positive influence through unthinking adherence to existing conventions. (Boud, 1998)

For the purpose of this study, it will be assumed that the method or assessment task adequately reflects the curriculum and it is valid, and that it requires the assessor to make professional judgements about the work of the learner. The issue then is how can the transparency, reliability and consistency of the task assessment process be assured. To achieve these goals, some forms of moderation or quality control and assurance must be implemented within the task assessment process.

Unreliability in the task assessment process can be due to inconsistency of individual assessors (poor intra-marker reliability) or inconsistencies across assessors (poor inter-marker reliability). Thus, the fewer the assessors, the easier it is to control the reliability factor. However, even with one marker, strategies need to be developed to ensure reliability. Although the literature discusses a number of strategies to improve reliability, such as double marking (S. Brown & Knight, 1994), or using a list of criteria (A. Miller et al., 1998), these strategies are often poorly implemented or not at all in universities (Boud, 1995, 1995a; Fry et al., 1999; O'Donovan et al., 2004; Webster et al., 2000). Boud (1995) notes:
There is probably more bad practice and ignorance of significant issues in the area of assessment than in any other aspect of higher education. This would not be so bad if it were not for the fact that the effects of bad practice are far more potent than they are for any aspect of teaching. (Boud, 1995a, p. 35)

Boud (1995) goes on to state that “assessors, be they staff or students, need the perspectives of others if they are not to be misled by the distractions of context and their own predictions” (p. 312), and this is what the moderation process is designed to provide. Moderation has been described as the “process of attempting to enhance reliability” (Gipps, 1994b, p. 12). Moderation processes are designed to improve the reliability of the educator’s professional judgement and may include:

1. Statistical moderation;
2. The use of moderators to compare standards of grading; and
3. Consensus meetings for educators to compare standards of grading.

Statistical moderation occurs when an external, usually quantitative assessment, is used to moderate the educators’ marking. This type of moderation may be very sophisticated, as with the application of Rasch modelling techniques (Athanasou & Lamprianou, 2002). Other forms of moderation comprise such methods as reviewing, re-grading, or independent grading of the students’ work usually by a moderator or another instructor (Council of Adult Education, 1994; Roberts & et al., 1996). To help assessors to adjust their marking to the standard, exemplars, occasional group meetings, or a moderator are used. Exemplars are samples of students’ work that exemplify a particular standard (O’Donovan et al., 2004). Two examples of the use of exemplars available online are: the American Exemplars website, which provides performance assessment tasks that meet the (American) national standards to improve assessment and instruction (Exemplars, n.d.); and the New Zealand English Online website (English Online, n.d.), which is part of an ongoing English professional development contract between Unitec Institute of Technology and the New Zealand Ministry of Education.

In primary and secondary schools, teachers mainly use clear statements of assessment criteria, a body of exemplars (or exemplification) of the outcomes, and informed professional discussion to increase intra-marker and inter-marker consistency. Meanwhile, in tertiary education, the taskforce on Assessment Policies and Practices (1996) set up by Queensland University, issued the following recommendations to improve professional judgement of educators in the assessment of student work.
For intra-marker consistency, the report recommended:

- Rigorous application of a clear statement of assessment criteria;
- Anonymous presentation of work by students;
- Staff development in marking;
- Adequate time frame; and
- Marking the same item for all students in sequence (as opposed to marking all items for each student in sequence).

While to maximise inter-marker consistency, the report recommended:

- Rigorous application of a clear statement of assessment criteria by all markers;
- "Conference" of markers to discuss standards;
- Double marking of papers receiving borderline marks;
- Group marking sessions so that all markers are subjected to the same conditions, and ambiguities are consistently resolved;
- Double marking of "problem" papers;
- One marker marking all the answers to each question; and
- Blind double marking of all papers, or of a random sample.

The above moderation processes are the generally accepted methods used to develop assessor familiarity with standards. They allow the assessor to decide what the appropriate moderated mark will be for a particular piece of work. However, due to reasons such as training, lack of time, money and resources, moderation is often carried out with many difficulties and does not meet its desired goal of consistency in marking.

The present use of print copies of exemplars is limited in number, not interactive and difficult to keep up to date; and the face-to-face professional discussions are rarely continuous, because they require a significant level of economic, time and personnel input to organise, run and manage. Thus assessors, due to the difficulties associated with traditional methods of moderation, often carry out assessment activities with poor professional knowledge of the required standards. As discussed earlier, involvement of moderation of the assessment instrument or marking key design and development could provide some assistance to overcome this lack of development of their professional knowledge.

The types of moderation referred to in the educational literature tend to focus on the marking process. While the goal of moderation is to improve reliability, validity and fairness of the marking, the possible ways that moderation might be applied to achieve
this are: developing the marking key collaboratively with the markers; moderating the samples before the start of marking; and meeting early into the marking period to fine-tune the standards.

**Assessment summary**

This section of the literature review identified assessment as an important and complex process within the field of teaching and learning that had major performance gaps between best or desired performance and current performance practice. While there is a growing trend towards alternative task assessment methods in the research literature, partly in response to pressure to improve authenticity, validity, reliability and transparency of the assessment process, the current practices have not kept up with the major changes that have occurred in other fields of tertiary education. The literature highlighted, among others, the need for improved transparency, feedback, reporting, and student involvement in the assessment process such as self and peer marking. Similarly, the process of quality control and assurance of the assessment process has not kept up with the major changes that have been and still are occurring in tertiary education.

While improvement in performance in any one area of assessment is desirable and has been demonstrated, the performance gains obtained tend to be time-consuming and not cost-effective. From the review of the existing literature on assessment, the researcher identified four areas within the task assessment process where performance could be improved. These areas were: 1) the marking key design, 2) the marking activities, 3) the skills and knowledge of the assessor, and 4) the management and administration process. In particular, the improvement in the management and administration of the task assessment process as a whole was identified as a key feature in the success of this study.

**Electronic performance support system**

This study proposed that the performance of the task assessment process could be improved through intervention strategies within the four areas identified in the previous section, and implemented through an electronic performance support system (EPSS). This section will introduce, discuss and highlight the literature on human performance improvement (HPI) with specific reference to EPSS.

EPSS falls within the field of HPI and is becoming the major focus of this field as
technology becomes ubiquitous within society. Human performance technology (HPT) is the methodology used by practitioners of HPI (Stolovitch & Keeps, 1999). The concern of HPT is with performance and return on investment issues and not training. The International Society for Performance Improvement (ISPI), whose mission is to ‘improve human performance in systematic and reproducible ways’, published an HPT model to support its mission (see Figure 2.4). The application of ICT and more specifically EPSS to the assessment process at present is limited and tends to focus on specific aspects. Gipps (2005) has suggested that “improvements in reliability and accuracy of marking, eradicate clerical errors, speed up marking process and, possibly, reduce cost” (p. 172) could be expected when ICT is applied to the assessment process.

**Human performance improvement**

In the field of HPI, a distinction is made between learning (training) or knowledge transfer that results in people knowing what to do but not necessarily doing it, and actual performance improvement (Fuller & Farrington, 1999; Pfeffer & Sutton, 1999; Rothwell, 2005). This study was interested in the latter, actual performance improvement of the task assessment process. Studies have suggested that there is often only a minimal permanent transfer effect of training of 10% to 30% unless other performance interventions are undertaken (Broad, 2005; Kirwan & Birchall, 2006; Stolovitch & Keeps, 2004). The present study was not interested in the assessor’s knowledge of marking, moderation and the task assessment process, but rather the implementation and application of this knowledge and skill to the task process. Learning and the learning technologies remain important to HPI, but it is clear that learning or training cannot merely be re-labelled as performance support, simply because there is a related performance issue. Unless there is an explicit performance outcome during the activity or task, it is not performance support or improvement (Dickelman, 1999).

HPI covers the areas of human activity, and the design and development of performance support systems (Stolovitch & Keeps, 1999). Sanders and Ruggles (2000) offer an excellent overview of the contributing disciplines that have influenced the development of the HPI field of study, including its beginning, back in the late 1940’s, with instructional design and programmed instruction. They state in their review that, “little information exists on exactly how particular performance consultants manage to get the organisational results they’re famous for” (p. 29). Gustafson (2000) confirms
this view of lack of documentation in relationship to the design and development of EPSSs.

HPT is a relatively new field of complex professional practice (Van Tiem, 2004), that has emerged to achieve the aims of HPI. Stolovitch and Keeps (1999), in their influential *Handbook of Human Performance Technology*, describe HPT as a *systematic* and *systemic* approach to improve the performance of individuals or groups. This approach to performance improvement seeks to eliminate, rather than overcome, the causes of poor performance. This is because elimination or removal tends to be less costly and time-consuming than the invention of new interventions.

The literature within the HPT field appears to use terms like performance support, job aides and performance improvement interchangeably. These terms are applied to systems, hardware, software or intellectual tools that enable the user to improve their effectiveness and efficiency when carrying out a job or task. Performance support systems are thus “environments that enable people to complete work with a minimum of training or learning in advance of doing a task” (Greenberg & Dickelman, 2000, p. 18), and where the cognitive load of memory and computation are reduced or eliminated (Norman, 1988) through the application of the support aid.

When ICT, especially in its digital form, is incorporated into HPI, the field of electronic performance support (EPS) is created. The literature on EPS has its beginnings in the field of HPT (Stolovitch, 2000; Stolovitch & Keeps, 1999). While this field is part of the dawn of the hoped for *paperless office*, the transition has been much longer than initially expected and is still a distant dream (Sellen & Harper, 2003). Thus, when one compares the affordance of paper with that of digital paper, paper still has many advantages such as being thin, light, flexible and portable. However, as access to ICT becomes widespread (ubiquitous), the affordance of digital paper is gaining on that of paper. The *paperless office* has gained a significant and growing role in the digital office and home. A recent example of this is a description of a *paperless* medical office by Gates and Urquhart (2007), where the authors note significant performance gains in cost and improved patient care due to reduced secretarial time in record keeping, improved tracking, and use of voice recognition software.

The literature abounds with definitions of EPSS, however a simple description is, where an EPS uses a computer software program to improve the performance of a work activity (B. Miller, 2001). Gery (1991) was a pioneer in the field and coined the term EPSS in the early 1990’s. The concept of support systems, of which EPSS is part, goes
beyond the boundaries of the individual person, to include enabling artefacts, environmental factors and interactions. The theory of distributed cognition (Greenberg & Dickelman, 2000) and the activity theory (Nardi, 1996) both provide theoretical foundations for an understanding and development of EPSS.

To design, develop and implement an EPSS, the HPT model is the one recommended in the field of HPI. The HPT model consists of five fundamental phases: 1) performance analysis, including a comparison between desired and actual performance, 2) cause analysis, 3) intervention selection and design, 4) intervention, implementation and change, and 5) evaluation. This model is exemplified and illustrated in Figure 2.4 (Van Tiem, Moseley, & Dessinger, 2004) and discussed in more detail in Chapter 3.

**Human Performance Technology (HPT) Model**

![HPT model diagram](image)

Gilbert (1996), one of the founding researchers in the field of HPT, identified six major areas that can contribute to workplace performance. He produced a model, now called the Gilbert Behaviour Engineering Model. Gilbert’s model shows these six areas divided into: environmental factors (information, resources and incentives) and individual factors (knowledge and skills, capacity and motivation). Wile (1996) combined five models of human performance, including this one of Gilbert’s, and
identified seven areas that he ranked according to “how often they are part of a performance problem” (p. 32) (see Figure 2.5). These areas highlight the full range and nature of performance factors that could be involved, used or modified to improve performance. This study focused on three of the performance areas identified by Wile (1996) as possible areas of intervention: Organisation Systems, Cognitive Support and Skills/Knowledge (see shaded boxes in Figure 2.5). Due to the limitations of the study, the areas selected were the only ones where intervention strategies could be implemented to improve performance of the task assessment process.

![Figure 2.5 Human Performance model. Source: Wile (1996).](image)

This study was interested in how a human computer-mediated environment, in the form of an EPSS, could support and improve the performance of staff involved in the task assessment process. The EPSS for this study endeavoured to improve performance through the integration and augmentation of the interventions strategies identified in the study from within the three intervention areas identified (organisational systems, cognitive support and skills/knowledge). The literature from the fields of distributed
cognition (Greenberg & Dickelman, 2000), action theory (Nardi, 1996) and HPI supports the development of a computer-mediated performance support system, based on the principle that individual cognition can be mediated by tools or aids and thus the cognitive load can be eased, reduced or eliminated through this mediation. The cognition (thinking process) thus becomes distributed and a social affair (see Vygotsky’s sociocultural theory (Daniels, 2005)). The implication of this is that “what a person can do with a tool is profoundly different than what a person can do without the tool” (Nardi, 1998, p 39).

The attempt at developing EPSSs for educators by Moore, Orey, and Hardy (2000) is one of the most recent and is still in the design phase. The skill and knowledge area of intervention, in this case the person setting the assessment and marking, would be captured and incorporated into the EPSS. This field of study is part of the knowledge management research literature and is discussed in the next section.

Knowledge management

Knowledge management (KM) “aims at integrating workplace knowledge and performance improvement” (Van Tiem, Moseley, & Dessinger, 2001, p. 27). KM is such a new field that a single definition has not been developed. A short comprehensive review of KM as it relates to performance management is provided in the chapter Performance Support Systems from the book Performance Improvement Interventions (Van Tiem et al., 2001). While the literature does not provide a concise definition of KM, it does involve capturing and making available to new or inexperienced users the intellectual capital of valuable professional knowledge and judgement of experienced users. In this study, this would involve the professional knowledge and experience of the course coordinator in relationship to the task assessment process. The performance aids used currently in the assessment process to achieve this are rudimentary marking keys and moderation. A detailed criterion or instructional rubric (see previous section) could be viewed as a method of capturing and distilling the objectives and goals of the assessment task that was set by the coordinator in the unit outline. This would also be available to future coordinators and other stakeholders of the unit and would be able to highlight in a much more transparent and detailed way the focus of the unit than the unit outline and the previous non-rubric marking keys.

There is a growing volume of literature that is promoting the combining of KM and HPI (Dickelman, 1999; Raybould, 2000; Schwen, Kalman, & Hara, 1998; Tillema,
40 Chapter Two: Literature Review

2005). Just making knowledge (e.g., e-exemplars) available even electronically is not sufficient; the knowledge needs to be made useful through a performance-centred interface, for example an instructional rubric, that incorporates the knowledge of the task setter, that is, their intellectual capital. The benefits of combining KM and HPI include the reduction of loss of knowledge, effective knowledge acquisition, and rapid knowledge access. This has become a focus for many organisations today, as Dickelman notes:

*Over the past several years organizations have been striving to make documents and other knowledge assets universally available. Those very assets form the basis of performance-enabling content if delivered in the right context. Further, Internet technology appears to enable ubiquity of knowledge assets like no other time in history.* (Dickelman, 1999)

Thus, the bringing together of KM and EPSS within the field of assessment offers the opportunity and possibility of significant gains in performance. These fields have grown and developed within the business world. In the field of education, computer-assisted assessment (CAA) has been developing, however until now the focus has been as a performance aid to the objective-based assessment. This study aims to expand this limited view of CAA to include subjective-based assessment. The following section briefly describes the current state of CAA.

**Computer-assisted assessment**

The discussion of EPSS would not be complete without the inclusion of a review of the literature on CAA. This is a broad term encompassing a range of applications, from the use of computers to conduct the whole assessment process, that is, on-screen testing (e.g., multiple choice questions), to only assisting in one aspect of the task assessment process (e.g., the optical mark and character readers) (Bull & Sharp, 2000). Bull and Sharp (2000) found that the computerisation of assessment has many advantages for both the assessment process, and assessors and lecturers. Assessment benefits were faster marking and feedback, and better student learning. They also claim that CAA enables assessors to “verify exam results without tedious manual assessment; and students receive immediate grade feedback. (With paper and pencil they have to wait to hear from their instructor, this might take one week.)” (p. 257).

Although CAA is an appropriate term, many authors have used a range of other related terms such as computer automated assessment (CAA), e-assessment (EA), online assessment (OLA), computer based assessment (CBT) and technology based
assessment (TBA) (Govender, 2003; Thomson Prometric, 2005). CAA is now being enhanced with the integration of the Internet (e-assessment) into a process of assessment that has major advantages, such as platform-independence and anywhere anytime access (Baillie-de Byl, 2004; Woit & Mason, 2003). The potential benefits of this type of product may be (Assessment and Qualifications Alliance, n.d.) (Pearson Education Australia, n.d.):

- Better quality marking, through early detection and remediation of aberrant marking;
- Random distribution of scripts and items to markers;
- Specialisation of markers in a limited number of items;
- Reduction of clerical errors, because the computer sums the marks;
- Elimination of paper distribution; and
- Greater security.

The first area of the task assessment process that took advantage of CAA was objective-based assessments that automated the marking process (eliminating the marker) and allowed the results to be instantly available. There are now many products on the market offering this type of assessment. For example, Question Mark Computing (Questionmark, n.d.) and WebMCQTM (MCQ International, n.d.), are at the forefront of the move towards computerisation of tests and assessments. These products provide services enabling the creation and presentation of interactive questions via the Internet. They are at present ideal for objective-based assessment, such as revision quizzes, formative exams, training packages and questionnaires, but lack the ability to assess subjective-based tasks. Meanwhile, automated essay marking is still on the horizon for everyday use and is currently being evaluated (Bampton, 2004; Christie, 2003).

The Computer-Assisted Assessment Centre (CAA Centre Resources, n.d.) carried out a major survey in 1999. This survey involved over 750 United Kingdom (UK) tertiary education staff, and found that the following factors, in order of importance, were critical for a successful use of CAA (Bull & Sharp, 2000, p. 257):

- Pedagogical and technical support;
- Time factors – savings in marking time, however, more time needed to develop suitable assessments;
- Confidence in the system, especially pedagogical fitness for purpose;
- Attitude of practitioners;
- Ease of use; and
CAA tools that enhance the feedback and reporting aspects of the assessment process are beginning to be developed. For example, the Australian product MarkIt (MarkIt, n.d.); and the Assessment and Qualifications Alliance (AQA), one of the largest unitary awarding bodies in the UK, that has recently started to use e-assessment and e-marking (Assessment and Qualifications Alliance, n.d.). However, the application of CAA to subjective-based methods of assessment is still being tested and developed. Thus, the importance and relevance of this investigate study into the design, development and evaluation of an EPSS to improve the performance of educators as they carry out the task assessment process. This study goes beyond the objective-based assessment tools that are currently available to assessors wishing to use CAA applications, and explores the application of an EPSS to subjective-based assessment.

Summary of review on EPSS

This section reviewed the literature on EPSS as it relates to the features and educational possibilities of such a tool to improve the performance of the task assessment process. Business has demonstrated significant performances gains are possible through the application of EPSS and KM. This study investigated whether these potential performance gains could be similarly achieved in the task assessment process where professional judgement is involved (subjective-based assessment). This section also covered the larger field of HPT, of which EPSS is part, and described the comprehensive HPT model developed by ISPI. This model is used as a map in the HPI field to improve performance and highlights the methods and steps needed to successfully implement improved performance in the workplace.

Design and development of an EPSS

This section of the literature review relates to phases 3 and 4 of the HPT model (refer to Figure 2.4), that is, the design and development, and implementation and change, as they relate to the present study. The literature on EPSS has been mainly centred on the commercial and not the educational world and, as suggested by Gustafson (2000) “there is very little literature available that describes how people have actually designed and developed EPSSs” (p. 42). Gustafson further states, “EPSS design is an immature technology about which there is much more to be learnt” (p. 42).
However, in the area of instructional design methodology, he does offer guidance to would-be designers in the use of task, audience and environmental analysis, and the use of formative evaluation tools and techniques. In the area of prototype development, he suggests the use of rapid prototyping, where a prototype goes through a number of iterations before the final prototype is accepted for release.

Raybould (2000) offers an alternative view to designing an EPSS, called ‘Performance Support Mapping Methodology’, which comprises elements of information engineering, business process reengineering, instructional systems design (usage-centred design and contextual design), knowledge engineering, and structured documentation. The key phases in this process are: look and listen, understand the work, design the work and design the interface.

These models are all bottom-up or outside-in approaches to design. Carr (1997) call these methods of design User-Design Power Dynamics, and Vredenburg, Isensee and Righi (2002) describe them as UCD. This is where the design, development and implementation processes “extends stakeholder involvement beyond mere input to create empowered users who have design and decision-making powers” (A. Carr, 1997, p. 6). Figure 2.6 highlights the complex power dynamics involved in this approach and the types of methods used in this UCD process. In this approach, when applied to this study, the designer (researcher) becomes the design facilitator and not the expert, the users are the tutors, and the leader is the coordinator of the unit.

![Figure 2.6 User-Design Power Dynamics. Source: A. Carr (1997, p. 8).](image)

The iterative or cyclic development process typical of action research is called prototyping in the commercial field of product development. The reason to use the
prototype method of development is in part due to Pareto’s Law (Figure 2.7). The law states that only a small amount of effort is required to produce a product that has most of the desired features and performance; however, a large amount of effort is needed to bring the product to completion. This law is often called the 80/20 rule.

![Figure 2.7 The curve of Pareto’s Law. Source: Nielsen and Mack (1994b).](image)

Isensee and Rudd (1996) identified eleven major advantages of prototyping, most of which are relevant to this study. These are: better collection of customer requirements; cost saving; increased quality; evaluation of new interface techniques and functions; early testing; and a better design. The scenario technique proposed by Nielsen (1993; 1994a) is a quick, cheap, valid and fairly reliable method of prototyping. Due to the scope and limitations of the study, this method was considered appropriate for consideration. It is now described in detail.

The scenario method developed by Nielsen (1994a) is a minimalist prototyping method (Figure 2.8), in that it describes a single interactive session without any flexibility for the user.

![Figure 2.8 Minimalist prototyping method. Source: Nielsen (1994a, p. 250).](image)

This method combines the limitations of both horizontal (low-fidelity) prototypes (users cannot interact with real data) and vertical (high-fidelity) prototypes (users cannot move freely through the system, however real data is involved). As a result, the
number of features and functionality are markedly reduced in this type of prototyping. Nielsen explains this approach to prototyping thus:

*The entire idea behind prototyping is to cut down on the complexity of implementation by eliminating parts of the full system. Horizontal prototypes reduce the level of functionality and result in a user interface surface layer, while vertical prototypes reduce the number of features and implement the full functionality of those chosen (i.e. we get to play with a part of the system).* (Nielsen, 1994a, p. 94)

Nielsen (1994a) recommends using three to five evaluators in the development of the prototype, because he believes that using larger number does not gain that much additional information.

This scenarios technique was based on Nielsen’s experience in the field of usability engineering and is part of the method he called the *discount usability engineering* method. This method incorporates a further two techniques: simplified thinking aloud and heuristic evaluation, both with an early focus on users. These techniques reduce the cost of design, development and implementation of the software. Nielson (1994a) reduced *thinking aloud* technique to its basics: some real users, and some topical test tasks (scenarios). Users are asked to think aloud while they perform the task, with the experimenter taking observation notes. Thus, this method is a way of getting quick and frequent feedback from users. Scenarios can be implemented as paper mock-ups or in simple prototyping environments.

The heuristic evaluation involves having individual members of the focus group examine the interface and judge its compliance with recognised usability principles. When all these evaluations have been completed, the members are permitted to communicate and have their findings aggregated. This procedure is important in order to ensure independent and unbiased evaluations from each member. In this study, the different views of the marking key were the interfaces under investigation and the users and focus group were the tutors and coordinator of the unit. A number of advantages are claimed for the heuristic evaluation method including (Nielsen, 1994a): low cost; intuitive to perform, since it is relatively easy to learn; and no advance planning required, since a single evaluator can carry out the evaluation.

There are two major reasons suggested by Nielsen for alternating between heuristic evaluation and user prototype testing. Firstly, a heuristic evaluation can identify and eliminate a number of usability problems, without the need to involve users, whose time and access are valuable and thus need to be used effectively.
Secondly, these usability assessment methods have been shown to find fairly distinct sets of usability problems that supplement each other and do not lead to repetitive findings. In the field of educational research the use of more than one method of data collection is called triangulation.

Nielsen (1994a) suggested that a small sample size, between three and five is not a problem in this type of usability prototype study. This small sample size not only simplifies testing but also has benefits similar to those of more complex testing and larger sample sizes. The curves in Figure 2.9 shows that the benefits of any user testing are much greater than the costs no matter the sample size; however, the maximum benefit-cost ratio is achieved with three to five users (Nielsen, 1994a).

![Figure 2.9 Cost-benefit trade-off curve. Source: Nielsen (1994a, p. 252).](image)

**Characteristics of the proposed EPSS**

Gustafson (2000) suggests that there are at least seven different design considerations when designing and developing an EPSS. For the convenience of the present study, each of these design factors have been listed on a continuum although some are not mutually exclusive. Table 2.3 lists these design parameters. Each will be discussed separately as it relates to this study.

<table>
<thead>
<tr>
<th>Black box</th>
<th>or</th>
<th>Glass box object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part-task</td>
<td>or</td>
<td>Whole task support</td>
</tr>
<tr>
<td>Embedded</td>
<td>or</td>
<td>Linked / external connect</td>
</tr>
<tr>
<td>Self-contained</td>
<td>or</td>
<td>Networked and shared work space</td>
</tr>
<tr>
<td>User controlled</td>
<td>or</td>
<td>System controlled</td>
</tr>
<tr>
<td>User modified</td>
<td>or</td>
<td>Organisation modified</td>
</tr>
<tr>
<td>Static</td>
<td>or</td>
<td>Dynamic system</td>
</tr>
</tbody>
</table>

**Table 2.3**

*Range of design parameters considered in the development of the EPSS*
The initial design concept and objective of the researcher in developing the EPSS was to make the task of assessment more efficient, without initially making the user more competent. Thus, the design initially focused on the black box concept. However, this concept needed to be monitored during the study to ascertain whether the involvement in the design and quality assurance of the marking key did contribute to the users’ competency.

The proposed EPSS would cover all activities involved in the task assessment process (whole task support) and be embedded in them and be self-contained and under user control, while the cognitive and clerical and administrative load would be reduced or eliminated. The user would have some degree of modifiability of the content and the system would be mainly static but have some dynamic features. As a result, the proposed EPSS should have a number of benefits and performance improvements:

- Improved consistency through the improved design of the marking key and the limitation of the range of possible grades or marks;
- Eliminated or reduced clerical and administrative activities;
- Improved feedback through the application of an instructional rubric and access to well thought out and considered feedback that could be applied when appropriate; and
- Availability of multiple views of the same data without requiring multiple data entry.

**Summary**

The literature reviewed in this chapter covered the three diverse fields of assessment, EPSS, and software product design and development. Assessment was shown to be complex and an important aspect of tertiary education that often drives the teaching and learning processes. There is a trend in Australian universities towards more learner-centred assessment methods that are more subjective and thus involve professional judgement on the part of the assessors. The current methods and processes of task assessment involving professional judgement tend to be based on tradition, impressionistic (global) marking, and based on a connoisseur model of assessment. This review of the literature identified a need to counter-balance this subjectivity and lack of transparency, and a need to improve the educative aspect of the task assessment process through improvements in: the design of marking keys; moderation; marking activities; feedback; constructive alignment of the assessment to the desired learning outcomes; and quality assurance of the whole assessment process.
The literature on EPSS supported the notion that a combination of KM and CAA could aid and support the performance of assessors as they carry out the task assessment process involving professional judgement. The proposed EPSS would include the improvements to assessment identified from the literature and would be developed through a user-centered, participatory, iterative, prototype design process recommended in the literature on software design and development.

Based on the literature review, the theoretical framework for the study began to emerge. This theoretical framework is described in the following chapter (Chapter 3). To match the three fields of research literature required to position the study, the theoretical framework required three distinct frameworks to fully conceptualise the study. These were: the overall investigation method framework, the area of investigation framework, and the design approach framework.
CHAPTER THREE
THEORETICAL FRAMEWORK

Introduction

This chapter discusses the theoretical and conceptual frameworks that were developed out of the literature review to inform the design of the study to address the research question: to what extent could the task assessment process that involved professional judgement be improved and supported through the application of an EPSS. In this context, the concepts that needed to be developed and explained were the overall method of investigation, the performance or task, and the method used to develop the performance interventions. These three concepts were conceptualised into three frameworks: the overarching theoretical framework of the study; the task assessment process framework (the performance activity); and the design and development method framework. These three interconnected frameworks will be referred to respectively as: 1) the human performance technology (HPT) framework, 2) the task assessment framework, and 3) the performance-centered design (PCD) framework. These frameworks developed from the literature review and emanated from three fields of literature respectively: the human performance improvement (HPI) field, the assessment field, and the software design field.

Firstly, the HPT framework evolved from a review of the literature in the HPI field, with specific reference to electronic performance support systems (EPSS) and intervention strategies. Elements of the HPT model developed by the International Society for Performance Improvement (ISPI) were chosen as the most appropriate on which to build and base the overarching theoretical framework of the study. This ISPI model provides a map that represents an all-encompassing approach to performance improvement in the workplace, from performance analysis to evaluation. The HPT model begins with a detailed investigation of the workplace activity and environment under investigation to identify a performance gap. For the purpose of this study, the workplace activity is the task assessment process, and the environment is the workplace.

Secondly, the task assessment framework developed out of the review of the literature in the field of assessment, with specific reference to the task assessment process, that is, the job or performance under investigation. The investigation of the
workplace activity, in this case the task assessment process, required its conceptualisation and definition. This proved difficult, as the researcher was unable to find any detailed or specific literature on the particular workplace activity that described the task assessment process in any precise detail. This meant that the conceptual framework of the task assessment process (the workplace activity) was general and limited in nature. However, the framework did inform and describe the task assessment process, and included the desired performance outcomes and the areas of possible performance intervention. The nature of the study meant that the workplace environment was the only possible performance intervention area, with only three of the possible seven areas of intervention identified by Wile (1996) within the scope of the study.

Thirdly, the PCD framework grew out of the literature review in the field of software design and development, with specific reference to performance or participatory design (PD) methodologies involved in the successful development of useful and productive software to improve performance. This framework describes how the interventions would be designed, implemented and evaluated. As the major intervention strategy was to be based around an EPSS, the PCD (Gery, 1991) and the user-centered design (UCD) (A. Carr, 1997; Vredenburg et al., 2002) were selected as the most appropriate.

Before proceeding, the researcher wishes to point out the issue of the inconsistent use of terminology between, and even within, these fields of research. During the review of the literature in the three fields, the researcher found that the terminology was often used inconsistently, as was noted, for example, with the use of assessment terms. However, limitations of space and time do not allow for the clarification of this issue and, if attempted, would have disrupted the readability of the thesis. The researcher therefore asks for the reader’s understanding that if at times they feel a term has been inappropriately used, the researcher has used it as it was cited in the literature.

The critical terms and concepts used in this study are: performance and performance improvement. These are described here as they are used in the study. Performance is defined as any kind of purposeful work or task. Meanwhile, performance support and improvement goes beyond mere knowledge of how to perform the process or task (Pfeffer & Sutton, 1999) and can involve a both external (tangible and intangible) and internal support (see Figure 3.2). This support can take many forms, such as: incentives, tools, aids, paper manuals, software, hardware, environment, and
organisational resources. Most of these can be considered or viewed as products, tools or more generally as systems.

Each of the three theoretical and conceptual frameworks described above are discussed in detail in the following three sections.

**The human performance technology framework**

Based on a review of the literature, the field of HPT was identified as the most appropriate from which to develop the overarching theoretical framework to be used in the study. The HPT model developed by ISPI was adapted, as represented diagrammatically in Figure 3.1, and adopted for this study. The researcher selected this HPT model because it is designed to be used as a “diagnostic and strategic tool for improving workplace performance” (Van Tiem et al., 2004, p 8), as previously discussed in the literature review (see Figure 2.6). The model focuses on performance improvement in a systematic and structured manner by the identification, design/development and implementation activities (interventions) to improve performance. In addition, the model describes these activities and suggests methods of evaluation.

The HPT model was modified by selecting elements that focused on the relevant aspects of the inquiry and minimise irrelevancies, as shown in Figure 3.1. This was essential due to the unique nature of the study, including that the prototyping of the interventions were performed live in the workplace without any pilot studies being carried out, and involved significant and critical live data that had to be error-free and performed within tight time constraints. This meant that the real world aspect of the investigation (i.e. the means by which validity in its various forms was achieved) was maximised. The users were also directly involved in the study and collaborated on the selection, design and evaluation of the interventions as recommended by the PCD model.
The limitation of resources available to carry out the study also meant that the HPT model needed to be modified and adapted. The resources available were limited to the researcher and the voluntary participation of the academic staff. Due to these limitations and unique characteristics of the study these five stages of the HPT framework were grouped into three distinct phases: firstly the Exploration Phase of the task assessment process, including performance and cause analysis; secondly the Development Phase involving the design, development and implementation of performance intervention strategies over a number of iterations; and finally the Evaluation Phase of these performance intervention strategies. Table 3.1 shows the relationship of the phases of this study to the HPT stages.

<table>
<thead>
<tr>
<th>Phases of Study</th>
<th>HPT Stages</th>
</tr>
</thead>
</table>
| Exploration Phase | 1 - Performance analysis (diagnosing workplace situations).  
|                   | 2 - Cause analysis (identifying causes).                                    |
| Development Phase | 3 - Intervention selection and design (selecting doable, adequate interventions).  
|                   | 4 - Intervention implementation and change (implementing changes).           |
| Evaluation Phase  | 5 - Evaluation (evaluating results).                                          |

The HPT framework shows the ways in which the stages interact and complement each other to inform the goal of performance improvement of the task assessment.
process. The early stages of the model tend to be linear, however the final three stages are iterative, interactive and non-linear due to the characteristics of the model.

The first two stages, *Performance Analysis* and *Cause Analysis*, establish the foundation for identifying the potential mix of interventions that could be employed to reduce the performance gap and thus improve performance. The first stage, *Performance Analysis*, consists of three types of analysis: organisational, environmental and performance gap analysis (see figure 3.1). These stages identify *what* is happening, that is, the actual or typical performance (current state) of the job or process, and the difference between this state and the desired (ideal) or exemplary performance requirements of the job. The difference between these two performances, the actual and the desired, is called the *gap*. Thus, a gap occurs when the two performances (desired and actual) do not align, or there is a performance problem.

The desired performance outcomes are derived from: the organisation’s mission, goals and policy statements, and experts and literature in the field under investigation. The desired performance outcomes of the task assessment process were initially identified from the research literature. These included a high level of validity, reliability, authenticity, transparency and fairness in the assessment process. Also identified from the literature was the requirement for quality feedback, reporting and management. These outcomes are discussed in the *Task assessment framework* section, and are further explored in the discussion of the *Exploration Phase* (Chapter 5) of the study.

The current or actual performance outcomes are derived from an analysis of the environments within which the performance occurs. This analysis involves both the external and internal environments within which the organisation operates. For this study, the external environment would constitute other universities, past and future students, future student employment organisations, and the wider community, including government and political bodies. However, this external environment was beyond the scope of the study. The internal environment consists of the organisation, i.e. the university, and all the levels down to the local workplace environment, including the staff, students and other internal stakeholders. Ideally, quantitative data would be available on the actual performance and this would be supplemented with qualitative information. However, within the literature, only very limited and rudimentary quantitative data exists on the task assessment process.
The *Cause Analysis* consists of a diagnostic analysis and is carried out to determine the *why* of the performance gap: what specific factors and areas caused the deficiencies, and what changes or interventions could be implemented to reduce the performance gap. These areas of performance interventions are described by Wile (1996) (see Figure 2.5). Wile’s model was modified for this study after a review of the assessment literature to highlight the areas where interventions for performance improvement of the task assessment might occur. This modified model is represented in Figure 3.2. A number of areas and sub-areas were identified as being outside or beyond the scope of the study, because they were either outside the local workplace environment or they were outside the control and influence of the coordinator, tutors or researcher. These areas were: financial incentives; the ICT infrastructure; and information, policies and guidelines provided by the university.

![Figure 3.2 Possible areas of intervention to improve performance. Adapted from Wile (1996).](image-url)

Based on the scope of the study and the review of the literature, the areas of possible intervention were identified as being: 1) the *organisational systems*, 2) the *cognitive support*, and 3) the *skills and knowledge* of the performer, in this case, the assessors (see shaded boxes in Figure 3.2). The *organisational systems* covered the activities involved in the task assessment process such as marking and management. *Cognitive support* or job aid is any resource or product that supports and enhances the
performance of the user. The final group of interventions within the scope of the study were the assessors’ skills and knowledge, such as on-the-job training.

The next two stages of the HPT framework (Figure 3.1), Intervention Selection and Design, and Intervention Implementation and Change, involve the selection, design and implementation of the best possible group of interventions. This selection is based on a consideration of appropriateness, cost, feasibility, and organisational and individual performer acceptability. To achieve maximum performance gains this selection needs to be in harmony and alignment with these considerations. These interventions are not implemented in isolation but are grouped, mixed and blended; and the acceptance by the users is crucial to the success of the intervention strategies applied. The methodology employed in this study to select, design and develop the interventions was a combination of participatory action research (PAR) and PCD approaches. PAR and PCD formed the bases on which the final framework, the PCD framework, was developed.

The Evaluation Stage, the final stage of the HPT framework (Figure 3.1), occurs throughout all the stages. This stage consists of formative, summative, confirmative and meta evaluation of both the processes and the products. Formative evaluation is used to diagnose and is on-going throughout all the stages. Summative evaluation focuses on the effectiveness of the performance intervention and occurs as many times as the intervention is modified and trialed. Confirmative evaluation occurs some time after the intervention has been implemented into the work environment, usually six to twelve months. The aim of this evaluation is to confirm the findings from the formative and summative evaluations. Finally, meta evaluation occurs after the final evaluation has been performed. In the HPT field, this type of evaluation is used to reflect on and evaluate the other three types of evaluation, and provide insight to the researcher or implementer.

The remaining two frameworks, task assessment and PCD, fit within this HPT framework and are described in the next two sections. The task assessment framework describes the task and work environment of task assessment as it relates to the study, while the PCD framework describes the design approach employed in the study to achieve performance improvement through the application of an EPSS.
The task assessment framework

The major components, areas and performance indicators of assessment at the workplace level were identified and analysed from a review of the literature including Edith Cowan University (ECU) policy and procedure documents. Based on this investigation, the literature in the areas of intervention from HPT (Figure 3.2), and the scope and resources of the study, the theoretical framework of the task assessment at the workplace level within the assessment process was conceptualised and is represented in Figure 3.3. This framework shows that assessment has two equally important processes: one undertaken by the learner (student), the assessment task, and the other undertaken by the tutor (assessor), the task assessment. As the focus of the study was on the performance support of the task assessment, the key performance areas or activities and areas of performance intervention were those relating to the task assessment (see shaded boxes in Figure 3.3).

The desired performance indicators of the task assessment at the workplace level are: the marking activities, the marking key, and the assessor skills and knowledge, including pedagogical content knowledge (Lederman & Gess-Newsome, 1999), and management and administration. These four areas represent potential opportunities where intervention and change could be introduced to improve the performance of the task assessment process. Marking covers such activities as: moderation; marking and grading of student tasks; the recording of feedback; and reporting, including the management and administration. The marking key represents the instrument used to assess the task, however the term has a wide range of meanings, from a guide to both students and assessor, to record feedback, or simply to record the marks and grades. The skills, knowledge and assessment literacy of the assessor was also identified as an important performance indicator. Finally, the management and administration area covers both the clerical work and the procedures involved in carrying out the task assessment process both at the assessment and unit level.

A number of intervention strategies at the workplace level were identified from the literature (refer to Figure 3.2) and were within the scope of the study. These intervention strategies were: organisational systems, cognitive support, and skills and knowledge. These potential intervention strategies represented opportunities to improve the performance of the task assessment process through the performance indicators. Organisational systems in this study refer to activities such as the marking and management activities, setting of clear goals, and job design. The cognitive support in
this study would be supplied through the EPSS and KM, and this support would contribute to the skills and knowledge of the workers.

Figure 3.3 The assessment process at the workplace level.
**Performance-centered design framework**

The third framework, PCD framework, developed out of the review of the literature, and informs the design and development of the interventions strategies associated with the four performance activities outlined in the previous section. As the focus and scope of the study was on performance support and improvement through interventions at the workplace level, specifically through an EPSS, the design focused on usability issues and aspects of the system. The term usability is used in this context to gauge or measure the success of a product. The International Standards Organization, (ISO) 9241-11, defines usability as (Bevan, 2001):

*The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*

Research has found that the performance of the support, be it a product, tool or more generally termed a system, is determined by both the usability of the support’s interface and how well it delivers precise amounts and type of support at precisely the correct time (Gery, 1991; Hackos & Redish, 1998; Norman, 1988, 1993). To achieve a high level of usability, a product usually requires user input into the design, development and evaluation. The greater the user input, the greater the likelihood of success of the usability of the interface and thus the performance provided by that support (A. Carr, 1997; Hackos & Redish, 1998; Nielsen, 1993; Vredenburg et al., 2002), while the closer this support is to the work or worker, the better. The involvement of users or workers in the design and development of a product, tool or system has been incorporated into a number of design approaches, such as UCD and PD.

The interaction and link between the user or performance and the system is called the interface. Hackos and Redish (1998, p. 5) described the interface as:

- The bridge between the user and the product or system and the world of the user;
- The means by which the user interacts with the product to achieve their goals; and
- The means by which the system reveals itself to the user and behaves in relation to the user’s needs.

Depending on the distance from the task or work, Gery (1991) divided this performance support into three categories: external; extrinsic or linked; and intrinsic or
embedded. External support consists of manuals, training and other resources not connected to the EPSS. Extrinsic or linked support is accessed from the EPSS interface but as separate entities, such as online help and cue cards. Intrinsic or embedded support is part of the EPSS interface, and includes icons, screens and wizards.

To achieve this improved system or product performance, the current literature within the design field continually emphasises the need for the worker (user) or the performance (activity) to be the focus of the design process and, where possible, for the user to be directly involved in the design process. The PCD design framework selected for this study, thus, was focused on user participation in the design and development process. This framework involved the active involvement of (workplace) practitioners in design and decision-making processes and shares many common values and techniques with PCD (Gery, 1991; Winslow & Bramer, 1994), UCD (A. Carr, 1997; Rubin, 1994; Vredenburg et al., 2002) and participatory action research (PAR). These design approaches focus on the user (performer) or performance, with the aim to improve the design and usability of a product or process. Meanwhile, the approach of the HPT model has a broader focus, and includes a whole range of interventions to achieve this improved performance support of a task.

These approaches to design were selected from the literature on software and product development as the most appropriate to be employed and integrated into the Development Phase of the study, and fit within stages 3 and 4 of the HPT framework. As the areas of interventions were located at the workplace level and involved the participation of the users, two key features of the PCD framework were well suited to this study: the involvement of users as equal partners in both the design and development of the product; and the use of an iterative prototype development process. The specific techniques and processes employed in this PCD framework are listed in Figure 3.4. They are based on the model developed by Nielsen (1994a) called the discount usability engineering method and are discussed in Chapter 2.
Summary

This chapter described the development of the three interconnected and related conceptual frameworks on which the method of investigation was developed. The HPT framework presented a well-developed and tested model for performance improvement covering: the identification of the performance issue or gap; the intervention identification; and the selection, design, development and evaluation. The HPT model was modified to accommodate the unique features of the study that will be discussed in detail in the next chapter. The significant areas and processes involved in task assessment were identified and described in the task assessment framework. The task assessment was the work activity under investigation and where the performance enhancements or interventions might be best carried out based on HPT methods. Theses possible areas of intervention and desired performance goals were identified based on the scope and limitations of the study. Finally, the framework for the selection, design and implementation of the intervention was described through the PCD framework.

The next chapter (Chapter 4) develops a methodology for the study based on the three conceptual frameworks described in this chapter. Chapter 4 firstly positions the study’s research approach and describes the type and nature of the overall research design, it then discusses each of the methods used within the three phases of the research design: the Exploration, Development and Evaluation Phases.
CHAPTER FOUR

METHODOLOGY

Introduction

This chapter discusses the research design used to investigate the research question that guided this study:

To what extent does the application of an electronic performance support system (EPSS) enhance and support the performance of the task assessment process where professional judgement is required to assess student work in a university course?

This chapter covers the following aspects of the research design: the development and description of the research design; the target population and setting; and the data collection and analysis.

The research design was based on the literature review, and further refined in the development of the theoretical framework that was discussed in the previous chapter (Chapter 3). All research projects, no matter how they are carried out, have in common that, “they all are, or aim to be planned, cautious, systematic and reliable ways of finding out or deepening understanding” (Blaxter, Hughes, & Tight, 2001, p. 5). Decisions about the position within a research paradigm and the choice of the research design and methods should be driven by the nature and aim of the research study and question(s) (Howe, 2003). The nature and aim of this study was to improve the performance and practice of the task assessment process, specifically when professional judgment is involved, through the application of an EPSS.

The study involved the design, development and implementation of an EPSS to aid, support and enhance the efficiency, effectiveness, satisfaction and efficacy of the task assessment process at the unit or workplace level. This EPSS would cover activities such as the marking key design, marking, moderation, feedback and reporting, and management. The focus of this study was on the impact this tool might have on the performance of the users involved in the task assessment process.

In educational research, strict positivistic experimental scientific research methods are less appropriate in that they tend to create a demand for divisions between the researcher and subjects (W. Carr, 1995). As this study was set in the educational research field, it involved complex issues and many variables over an extended period,
and was interested in practice and usefulness, such a scientific or experimental design method was considered inappropriate (McMillan & Wergin, 2002). As Kember (2003) states, “it is paradoxical that the more an experiment is controlled or shortened to enhance reliability and feasibility, the less relevant it is” (p. 96) to the application and evaluation of educational innovation such as that involved in this study. The author (2003) further discusses many of the issues that are associated with the application of the traditional or experimental design approach to educational research in tertiary education. Thus, based on the nature of the research question, the most appropriate philosophical orientation or research paradigm for this study needed to be naturalistic or interpretive in approach (McMillan & Wergin, 2002; Middlewood, Coleman, & Lumby, 1999).

Having thus placed the study in the naturalist paradigm, the type of research design approach was selected to inform the research question. While the idea of choice between distinct types of research, such as descriptive, explorative, predictive, explanatory, or action research, is in the end spurious, the real choice is a combination of the most valuable and useful types to achieve the aims of the study (Clough & Nutbrown, 2002; Howe, 2003). Vredenburg, Isensee and Righi (2002), in their book on user-centered design (UCD), also emphasise the importance of selecting a set of research methods that will provide the necessary information to answer the desired research questions. They identify three core types of UCD methods: “understanding users, evaluating design and performing hands-on testing for competitiveness” (p. 77) that they suggest need to be included in every project or study. Thus, the context and purpose of the research study should determine the choice of research design and methods to be used to address the research question and subsidiary questions.

This study considered a number of related research designs in the academic and commercial fields within the naturalistic or interpretive paradigm. The selection of the research design and a group of methods was driven by the conceptual framework, in conjunction with the research question. This process of analysis and selection resulted in the research design being positioned within the following research areas: applied, exploratory, participative action, evaluative and qualitative. This selection is summarised in Figure 4.1, based on a figure from Kumar (1999, p. 8) (shading indicates where this research design fits within the various research classifications).
Development of the research design

According to Clough and Nutbrown (2002), social and educational research “sets out with specific purposes from a particular position, and aims to persuade readers of the significance of its claims; these claims are always broadly political” (p. 4). The research design enabled this study to be persuasive, purposive, positional and political. As with all research, this study’s design and methods were based on a unique set of circumstances and determined in part by the scope and limitations of the study. This covered such aspects as: intervention strategies limited to the unit or workplace level; time constraints; infrastructure limitations; limited resources both financial and technical; situational and authentic environment; experience of team members; and use of high-fidelity prototypes in an actual work-environment instead of a simulated environment. The focus of the research design was on the design, development and implementation of a number of iterative high-fidelity prototypes of an EPSS set in an authentic environment. These were designed to improve practice within the task assessment process and thus help to inform the research question.

A number of different research designs within the field of product design, development and evaluation in both the research and the commercial domains were considered suitable for this study. These types of research designs have one or more of the following characteristics in common: action-based, problem-based, cyclic in nature, outside-in design, active user participation, responsive in the design and development
process, user-centered, and performance- or task-focused (Campbell, Gilroy, & McNamara, 2004; Cherry, 1999; Hinchliff, 2004; Middlewood et al., 1999; Murray & Lawrence, 2000; Nielsen, 1993; Vredenburg et al., 2002). These research design characteristics were considered essential to the successful design, development and implementation of the EPSS and to inform the research question. As stated previously, all studies are unique, and thus a combination and blending of a number of these research designs was necessary.

The research designs that were considered for this study and that included some or all of the aforementioned characteristics, focused on applied research with specific reference to action research. Applied research aims to “assess, describe, document or inform problems from the stakeholder’s prospective” (Anderson, 1998, p. 121) and to find a solution to them (Wiersma, 2000). While action research (Cherry, 1999; Dick, n.d.; Elliott, 1991) is a family of research methodologies that combine action (or change) and research (or understanding) at the same time. Pring (2000) describes the aim of action research as the improvement of practice as “contrasted to ‘normal’ research that aims to produce new knowledge” (p. 131). Pring further clarifies this difference by pointing out that “it is not enough to claim that practice has improved but that there is knowledge of why it improved” (p.134). Action research is usually a “cyclic or spiral process which alternates between action and critical reflection and in the later cycles, continuously refining methods, data and interpretation in the light of the understanding developed in the earlier cycles” (Dick, n.d.).

Participative research is a sub-group of action research that directly involves the participants in the research. This type of research has been called participative action research (PAR), participatory design (PD) or co-operative inquiry research in the academic field (Anderson, 1998; C. Marshall & Rossman, 1999; Whitmore, 1998; Wiersma, 2000). Participative research, in contrast to the scientific approach, places the emphasis on the collaboration between researcher and subject, to the extent that the distinction between them is removed, and all those involved act as co-developers or collaborators, contributing equally to the decisions that inform the research and to the actions of the research study. These participative research methodologies could be considered to fall within the UCD (Vredenburg et al., 2002) or PAR approaches that are used in the commercial field. However, the study did not set out to apply PAR.

Vredenburg, Isensee and Righi (2002) describe and introduce the current thinking on UCD methodology in their book *Use-Centered Design: an integrated approach.*
UCD is a type of action research that is used in the commercial world for the development of usable products and systems. A feature of this type of design is the focus on the user or an outside-in approach to design that should be:

... based on the needs and characteristics of the user audience - resulting in a usable tool - and on enabling efficient performance of their tasks - resulting in a tool that is useful as well. In other words, to design a productive human-computer system, the overall design of the computer system should proceed outside in - from the perspective of the users and their tasks - rather than inside out - from the perspective of architecture and code design (Scanion & Percival, 2002).

Vredenburg et al., (2002) emphasise the importance of selecting a core set of “UCD methods that will provide the information on user tasks and environment, their current ways of carrying out their tasks, iterative rapid feedback on low- and then high-fidelity prototypes, focused and comprehensive hands-on testing, post-shipping feedback” (p. 76). In UCD it is important that all “aspects of the product are customer driven” (Vredenburg et al., 2002, p. 2), from the features to the user interface, including the total customer experience. The customer or user expectations have changed over time, from the main focus being on the behaviour dimension of the product, to the cognitive and finally the affective dimension, with the trend in the future being focused on “delight – addressing all three dimensions simultaneously and in all the right amounts” (Vredenburg et al., 2002, p. 197) (see Figure 4.2). For the purpose of this study, the innovative product to be developed was the EPSS, the customers or users were the tutors (markers) and coordinator, and the goal of the product was the performance improvement and enhancement of the task assessment process.

![Figure 4.2 User expectation of a product. Source: Vredenburg et al. (2002, p. 179).](image-url)
Description of the research design

The aim of the study was the use of strategies incorporating an EPSS to support and enhance the task assessment at the workplace level where professional judgement is involved. These strategies and EPSS would support the development of the marker’s knowledge, marking key design, marking activities, feedback and reporting, and management of subjective assessment tasks. The EPSS aimed to be interactive, on demand and ubiquitous, and promote efficiency and effectiveness of the task assessment process. To achieve this aim the research design required three distinct phases, each involving a distinct research design. To implement this multi-phased research design the researcher used a modified and adapted Human Performance Technology (HPT) model as described in Chapter 3 (Figure 3.1) as the overarching framework that combined and linked the three phases together.

This HPT framework was used to give structure and shape to the research study, to reduce the complexity of the research design, and to aid in the justification and selection of the methods used. The study was divided into three phases based on the conceptual HPT framework. These three phases were the Exploration Phase, the Development Phase, and the Evaluation Phase. The Exploration Phase incorporated stages 1 and 2 of the HPT model, that is, the Performance and Cause Analysis. The Development Phase incorporated stages 3, 4 and 5 of the model, that is, the selection and design, implementation, and evaluation of the intervention strategies initially selected in the Exploration Phase to improve performance. The final phase, the Evaluation Phase, incorporated stage 3 of the model and covered formative, summative and confirmative evaluation. An overview of the phases of the study is shown in Figure 4.3.
Figure 4.3 Study design and phases.

The three phases of the study, *Exploration, Development and Evaluation*, were designed to answer different aspects of the research question, and required different methods of data collection in order to inform the research question. The relationship of these phases to the main research question themes drawn from the research question of the study are summarised in Table 4.1.
Table 4.1  
*Time frame and research questions as they relate to the phases of the study*

<table>
<thead>
<tr>
<th>Phase</th>
<th>Time (months)</th>
<th>Research Question Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploration Phase</td>
<td>3</td>
<td>Design Themes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are the possible key design features that contribute to the performance gap?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are the possible key design interventions that could reduce the performance gap?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overlaps with the next phase</td>
</tr>
<tr>
<td>Development Phase</td>
<td>12</td>
<td>Design Themes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are the possible key design interventions that could reduce the performance gap?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overlaps with the previous phase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Usability Themes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What components of the EPSS do users find useful?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What common patterns of use of the EPSS, if any, by the users are evident?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Implementation Themes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What constraints or obstacles exist to the effective use of such an EPSS?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What effect does the EPSS have on the moderation processes?</td>
</tr>
<tr>
<td>Evaluation Phase</td>
<td>1</td>
<td>Design Themes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are the key design features to reduce the performance gap?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overlaps with the previous two phases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Usability Themes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What components of the EPSS do users find useful?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What common patterns of use of the EPSS, if any, by the users are evident?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Implementation Themes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What constraints or obstacles exist to the effective use of such an EPSS?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What effect does the EPSS have on the moderation processes?</td>
</tr>
</tbody>
</table>

**Exploration Phase**

The first phase, the *Exploration Phase*, involved the performance and cause analysis of the task assessment process. This phase was designed to gather information about the gap between the desired and actual performance of the task assessment and the possible causes of these gaps. Due to the scope and limitations of the study, a rigorous analysis of the gap or causes between the actual and desired performance of the task assessment was not feasible or required. The aim of this phase was limited to giving guidance and direction to the *Development Phase*, that of the design, development and implementation of the EPSS.

Thus, the *Exploration Phase* involved a brief investigation of the activities involved in the task assessment, that is, the marking, moderation and management activities carried out by the tutors and coordinators. The desired performance was based on the current best practices identified from the literature in the task assessment process.
and from the university policy documents on assessment. The actual performance of the task assessment process was determined through observation and through information gathered on issues involved in the current marking, moderation and management of the task assessment process experienced by both tutors and coordinators at Edith Cowan University, with specific reference to the School of Education. The information gathered during the exploration phase was used to inform the initial intervention strategies and the design of the EPSS, highlighting the gaps between the actual and desired tasks and outcomes.

Development Phase

The second phase, the Development Phase, consisted of the design, development and implementation of selected intervention strategies embedded within an EPSS to improve and enhance the performance of the task assessment process. This approach represented a modification of the HPT model in that the study had preselected the use of the EPSS as the main means of implementation of the selected intervention strategies. The design approach was a bottom-up or outside-in approach, that Vredenburg, Isensee and Righi (2002) describe as UCD model and Carr (1997) calls ‘user-design Power Dynamics’ model. This modified model, illustrated in Figure 3.4 (Chapter 3), highlights the complex power dynamics involved in this type of approach and the methods used. This approach “extends stakeholder involvement beyond mere input to create empowered users who have design and decision-making powers” (A. Carr, 1997, p. 6). In this type of design approach the designer, in this case the researcher becomes the design facilitator and not the expert. For the purpose of the study, the users are the tutors, with the leader being the coordinator of the unit.

The Development Phase was the critical phase of the study and was carried out over two semesters in six cycles. This phase involved the combination of two research approaches: the PAR approach, where users of the product are collaborators, and the UCD approach, where users are directly involved in all aspect of product development. Both methodologies emphasise the cyclic or iterative design approach to the investigation of a problem or product development.

Evaluation Phase

The Evaluation Phase occurred after the Development Phase and was the last phase of the study. The methods of data collection are described in the Data Collection
As can be seen from the HPT model (Figure 3.1), evaluation occurred throughout the study and involved three types: formative and summative evaluation, associated with the Development Phase, and confirmative evaluation, that occurred sometime after the completion of the Development Phase. The formative evaluation occurred throughout the Development Phase during each of the development cycles (the findings of each development cycle are discussed in Chapters 6 and 7). The focus of the formative evaluation was on the usability of the intervention strategies incorporated into the EPSS to meet the task analysis gaps identified in the Exploration Phase and during the development cycles of the Development Phase.

The final confirmative evaluation, although not part of the original research design, was made possible because the coordinator involved in the Development Phase wanted to continue using the EPSS. In addition, other unit coordinators that had been following the study wanted to trial the product. Even though the EPSS was still in a beta version of development, the product was complete and stable enough to require minimal involvement of the researcher in the day-to-day use of the EPSS.

The number of iterative development cycles, six and the time span of one year that occurred in the Development Phase reduced markedly the halo or Hawthorne effect due to the researcher being present and enthusiasm of the team members being involved in something new and innovative. While the post Development Phase of the study reduced the involvement of the researcher, allowed for confirmative evaluation and triangulation to occur, and further reduced the halo or Hawthorne effect due to the researcher being not present and enthusiastic about the innovation.

**Target population and setting**

The subjects or users involved in the Exploration and Development phases were drawn from the academic staff (both full-time and sessional) in the School of Education at Edith Cowan University. Participation in the Development Phase had, however, further requirements: the participants had to be involved in units that had at least four tutors and one coordinator. This was the minimum number of evaluators suggested by Nielsen (1994a) for this type of usability prototype study. The units also had to have a significant part of the overall total mark derived from assessments that required assessors to make professional judgements. Thus, to meet these requirements of the Development Phase of the research project, only large education units were considered for selection.
Additionally, the participants in this phase had to be already familiar with the processes involved in moderation. This additional requirement was felt necessary after the *Exploration Phase*, when during informal interviews, some tutors who were inexperienced and/or without an academic background in education, indicated either that they did not understand or that they did not appreciate the need for reliability, consistency and moderation in the assessment process. To involve these tutors in this study would have meant that valuable time would have had to be allocated to professional development in this important area of assessment.

These limitations to the unit selection for the *Development Phase* were important and necessary to allow the study to focus on the assessment processes, as both tutors and coordinator would be already familiar with the needs and benefits of improving the marking key, moderation and management processes. As many of the tutors involved in the project would be sessional and would not be paid for attending regular team meetings, this prerequisite knowledge of assessment was essential for the success of the project.

From the contacts made during the *Exploratory Phase* and through a process of informal discussions with unit and programme coordinators in the School of Education, two first-year units were selected that met the study requirements and where both the coordinator and tutors were willing to participate. The unit selected for the first semester was EDL1101, *Learning and Development I*, while for the second semester the unit selected was EDL1200, *Learning and Development II*. Each of the units included three different assessment items requiring professional judgement, and involved four tutors and the coordinator. In the second semester, the researcher was fortunate to continue the collaboration with the same coordinator and two of the tutors from the first semester team.

This *Development Phase* of the study was thus set in the authentic university environment of two first-year pre-service teacher education units (courses) in the School of Education, at Edith Cowan University. This phase of the study took place over two semesters in 2003, and the units’ coordinator and tutors were directly involved as collaborators.

The context and the requirements of this phase of the study presented a number of design constraints and limitations. The main design constraints related to the study being set in an authentic environment: the assessment methods were pre-set and could not be modified; participants experienced the stress of assessing real student work; and
the researcher experienced the difficulty of working with sessional staff, who were only on campus usually around their tutorial and meeting times and often had other commitments. The collaborative nature of the iterative development process added to the challenges of this Development Phase.

These constraints were unavoidable and reduced flexibility, however, setting the study in an authentic environment was pivotal to the success of the study because it meant that the EPSS was designed, developed and implemented in a real and authentic working environment with real workers and assessment processes. This resulted in findings that had a higher ecological validity and reliability than if the EPSS had been developed using a more traditional and commercial design, develop and implementation cycle, that is, where the prototype is evaluated under simulated conditions. Thus, the study placed the team members including the researcher in a unique situation of prototyping in a real working environment. Further, this method of development also helped to overcome the problem that education research and policy are not often taken on board or valued by educational practitioners (Hayward & Hedge, 2005; Reeves, 1999), as practitioners themselves were involved in the design and development of the EPSS.

Data collection

The data collection methods were selected from both the academic and commercial fields previously discussed, and modified according to the specific requirements of the study. The data sources, methods of collection and purposes were different in each of the phases and are summarised in Table 4.2. They are described and explained in the order in which they were used in the three phases of the research design: the Exploration, Development and Evaluation Phases.

In the Exploration Phase, there were three sources of data: tutors, experienced coordinators, and policy documents and literature review. The purpose of this data collection was to set the scene and background from which to begin to design and develop intervention strategies to be incorporated into the EPSS. The second phase, the Development Phase, involved three sources of data collected over one year. The three sources of data were the researcher, unit team members and the team as a whole. Based on this data, the EPSS was refined and developed over six iterative development cycles. The final phase, the Evaluation Phase, occurred after the completion of the Development Phase, and consisted of summative and confirmative evaluation. The
source of data for the summative evaluation was the unit team members, while the data for the confirmative evaluation originated from the users who continued to use the EPSS after the study; and from presentations, journal articles and awards related to the study.

Table 4.2
Data sources, methods of collection and purpose by phase of the study

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Method of collection</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exploration Phase</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tutors</td>
<td>Structured Task</td>
<td>To understand what the tutor does during the task assessment process and the environment in which the tutor works (contextual inquiry).</td>
</tr>
<tr>
<td></td>
<td>Observation</td>
<td></td>
</tr>
<tr>
<td>Experienced coordinators</td>
<td>Semi-structured</td>
<td>To understand what experienced coordinators consider best practice in the task assessment process.</td>
</tr>
<tr>
<td></td>
<td>Interviews</td>
<td></td>
</tr>
<tr>
<td>Books, journals and e-sources</td>
<td>Literature Survey</td>
<td>To expand on the two previous data sources by carrying out a literature review.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECU and other university assessment</td>
<td>Document Analysis</td>
<td>To determine the desired performance goals of the task assessment process as mandated by the university administration.</td>
</tr>
<tr>
<td>documents</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Development Phase</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Formative evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Researcher</td>
<td>Minimalist Prototype</td>
<td>To reduce the number of features and functionality and thus complexity and cost of design and implementation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Team Members</td>
<td>Structured Task</td>
<td>To observe how, what aspects of, and when the EPSS is used.</td>
</tr>
<tr>
<td></td>
<td>Observation</td>
<td></td>
</tr>
<tr>
<td>The Team</td>
<td>Heuristic Evaluation</td>
<td>To examine the marking key content, judge its compliance with assessment objectives, educational goals, interface and recognised usability principles.</td>
</tr>
<tr>
<td></td>
<td>and Focus Group</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td></td>
</tr>
<tr>
<td><strong>Evaluation Phase</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Summative evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit Team Members</td>
<td>Semi-structured</td>
<td>To inform the research question and theme questions. To gather information from the tutors about: usability, suggestions, difficulties, attitudes, feelings and reactions to using the EPSS in the assessment process.</td>
</tr>
<tr>
<td></td>
<td>Interviews</td>
<td></td>
</tr>
<tr>
<td><strong>Confirmative evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tutors &amp; coordinators using EPSS</td>
<td>Semi-structured</td>
<td>To inform the research question and theme questions. To gather information from the coordinators about: usability, suggestions, difficulties, attitudes, feelings and reactions to using the EPSS in the Post-Development Phase.</td>
</tr>
<tr>
<td>after study</td>
<td>Interviews</td>
<td></td>
</tr>
<tr>
<td>Conference presentations,</td>
<td>Documentation of</td>
<td>These documents, presentations and awards show that the study has gained recognition from the wider academic community.</td>
</tr>
<tr>
<td>journal articles, awards</td>
<td>Material</td>
<td></td>
</tr>
</tbody>
</table>

Data collection during the Exploration Phase

The *Exploration Phase* was limited in scope and time, and involved a brief task analysis of the task assessment process. This was done to identify and describe the gap
between what was actually done by the tutor during the task assessment process and the desired goals of the task assessment. The scope was limited as the focus of the study was on the design and development of the EPSS, while the time was limited to only three months due to the overall time constraints of the study. Furthermore, the Exploration Phase was limited because the findings were only used to guide and help the first development cycle of the Development Phase.

The Exploration Phase involved the collection of data from three sources that required the application of three different collection methods: structured task observation of tutors and coordinators during the marking process; best practice experts interviews of experienced Edith Cowan University coordinators; and a literature and policy document investigation.

Firstly, the structured task observation method was used to identify and describe current workflow practices that were being used in the task assessment process. This involved the description of the task assessment performance of the tutors as they were carrying out the task assessment process. The observation focused on how the tutors were actually marking and how they were using technology in the process.

Secondly, to obtain information on the desired goals and current best practice involved in the task assessment process, a number of semi-structured interviews were carried out with experienced coordinators and experts in the field of assessment at Edith Cowan University.

Thirdly, a literature and policy document review was used to investigate both the current and the desired practice, and allow triangulation to occur.

The findings from this phase were used to help set the scene for the study and helped guide the initial design parameters of the EPSS that were used in the Development Phase.

Data collection during the Development Phase

The Development Phase formed the major part of the study and helped to answer the design and usability themes of the research question; this phase took one year to complete. The methods used in the design, development and implementation phase to inform the research question were based on the participatory action research and UCD methodologies discussed in the previous chapter (Chapter 3), and they were applied to six assessment cycles.
The data collection methods selected for this phase are described in Table 4.2. These methods were: firstly, the development of a minimalist prototype, whose purpose was to reduce the number of features and the functionality of the EPSS, thus minimising the cost of design and implementation. Secondly, a structured task observation of tutors carrying out the task assessment using the EPSS, to identify how, when and where the EPSS was used by tutors and coordinator during the task assessment process. Thirdly, a heuristic evaluation was carried out through focus group discussions conducted during normal meeting times and involving the team as a whole. The purpose of this heuristic evaluation was to assess the compliance of the EPSS with assessment objectives, educational roles, and usability principles (heuristics).

These research methods were based on the UCD and PCD methods (Nielsen, 1993, 1994a; Nielsen & Mack, 1994b; Vredenburg et al., 2002; Whitmore, 1998) and were considered appropriate for the Development Phase, where an iterative development and formative evaluation process was necessary for the design, development and implementation of the EPSS. In evaluating designs, Vredenburg, Isensee and Righi (2002), suggested that the focus should be on “gathering user feedback to the evolving design” (p. 34) and that the user feedback should be gathered often, with rigor and speed, and drive product design.

The choice of research methods in this phase was in part due to the importance the research literature placed on user participation and their commitment to the success of this type of innovative product development. In addition, because the study required responsiveness, flexibility, action and change to be carried out within short iteration cycles of between four to six weeks, participant active involvement was essential. These selected methods allowed the researcher to continuously test the underlying assumptions of the EPSS rigorously. They also enabled the tutors (practitioners) to collaborate and participate actively in the iterative process and thus take ownership of both the changes and the EPSS.

During the Development Phase, data were gathered through six iterative development cycles (see Figure 4.4). Each development cycle was brief, taking usually only four to six weeks to complete, and often overlapped with the following cycle. Team meetings occurred usually on a weekly basis, and tutors were observed regularly using the EPSS during each development cycle and as the need arose. Modifications to the EPSS were based on coordinator’s and tutors’ observations and feedback, the data obtained through the structured task observation, and the researcher’s own reflection on
all activities of each development cycle. Relevant literature when and where appropriate was provided to help both coordinator and tutors understand the theory and practice behind the rubric method of marking key design and marking. Thus, the coordinator and tutors were collaborators and were able to directly interact with the researcher with comments, suggestion and ideas. This interaction improved over time and the team formed a strong bond that has continued after the completion of the study.

![Figure 4.4 The task assessment development cycle.](image)

**Data collection during the Evaluation Phase**

The *Evaluation Phase* occurred at the completion of the *Development Phase*. This phase included two distinct types of evaluation: summative and confirmative. The summative evaluation occurred soon after the completion of the *Development Phase* and involved the use a semi-structured interview technique. The purpose of the summative evaluation was to collect data from individual team members about their experiences and observations in using an EPSS to complete the task assessment process. The confirmative evaluation was on-going following the completion of the *Development Phase*, and involved semi-structured interviews of users of the developed EPSS prototype conducted around a year after the completion of the *Development Phase* and on-going academic acknowledgment of the study through presentations at conferences, journal articles and awards.
Data analysis

An inductive approach was applied to the data analysis and, while specific questions were posed to guide the study, no hypotheses were formed prior to data collection and analysis. Wherever possible throughout the study, an effort was made to gather multiple perspectives through multiple sampling points to enable triangulation of the data gathered (P. Marshall, 1997) and cross-checking to improve reliability. A summary of the links between the research themes and questions, phases of the study and the different data collection methods used is shown in Table 4.3.

Table 4.3

<table>
<thead>
<tr>
<th>Themes and questions</th>
<th>Phase</th>
<th>Data Collection Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design Themes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What are the key design features of an on-screen marking system?</td>
<td>Exploratory Phase</td>
<td>Semi-structured interviews of expert group</td>
</tr>
<tr>
<td><strong>Usability Themes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How is marking currently done with reference to ICT usage?</td>
<td>Exploratory Phase</td>
<td>Observation of markers using the EPSS</td>
</tr>
<tr>
<td>What components of the on-screen marking system do users find useful?</td>
<td>Development Phase</td>
<td>Observation of markers using the tool</td>
</tr>
<tr>
<td>What common patterns of use of the on-screen marking system by users are evident?</td>
<td>Development Phase</td>
<td>Observation of markers using the tool</td>
</tr>
<tr>
<td>What constraints or obstacles exist to the effective use of such an on-screen marking system?</td>
<td>Development Phase</td>
<td>Observation of markers using the tool</td>
</tr>
<tr>
<td>What effect does the on-screen marking system have on the moderation processes?</td>
<td>Development Phase</td>
<td>Observation of markers using the tool</td>
</tr>
</tbody>
</table>

The methods used in data analysis were all qualitative due to the nature and scope of the study. The data were analysed by looking for themes and trends, and it was also used for descriptive purposes. The themes used to analyse the data were derived from the areas of significance that were identified and developed from the literature review and the theoretical framework covered in the last two chapters (Chapters 2 and 3). They formed part of the results of the Exploration Phase and involved the performance, gap and cause analysis. These theme areas were: marking key, marking activities, skills and knowledge of assessors, and management.
An extensive organisational/institutional and literature review of policy and procedures was carried out in the area of assessment. This documentation was analysed to identify the desired performance outcomes of the task assessment process. To further determine the desired performance a number of expert assessors were interviewed using as semi-structured method. Due to the small number of interviews, the reliability was necessarily low but the validity was high as they were all experts in the field. As this phase of the study was only explorative in nature, this low reliability was acceptable.

The observation of tutors occurred during semi-structured interviews. This observation occurred in brief snapshots as the tutors carried out the task assessment process within the normal work environment. The analysis of the observation data was designed to determine the current or actual performance of the task assessment within the environment being studied, and the structure used to record and group the observation data was based on the themes already developed. The reliability was low due to the small number of observations but this was allowed for in the study, as this phase (Exploration Phase) was only explorative in nature. The validity was high, as the observations occurred live within the normal working environment.

The observation of tutors during the Development Phase using the EPSS occurred frequently, but informally and in an ad hoc manner. The researcher worked with the tutors and coordinators explaining, helping, encouraging, and solving EPSS technical issues as they occurred during the development cycles. Validity and reliability were high as these observations occurred on a regular basis and throughout the year of the Development Phase. The results from the observations and feedback were usually immediately incorporated into the task assessment strategy and EPSS or into the next development cycle.

The Evaluation Phase involved both summative and confirmative evaluation. The summative evaluation involved the semi-structured interviews of the tutors and coordinator and analysis of this data. This was carried out at the completion of the Development Phase. Validity and reliability was high as the interviews were carried out individually over one hour, and used a semi-structured approach using the themes developed and refined during the Development Phase. This approach helped to guide, focus and keep the interview on track and helped in the data analysis. The confirmative evaluation occurred sometime after the study was completed and involved semi-structures interviews of users of the EPSS, and the analysis of awards, documents and
conference presentations that show the study has gained recognition from the wider academic community.

**Ethical considerations**

The research methods were approved by the Edith Cowan University Ethics Committee before the commencement of the study. This was to ensure the rights of the participants would be protected, and their anonymity guaranteed. Written permission, containing description of study and right of withdrawal were obtained from all participants. All data collected has been stored securely at Edith Cowan University and will be held for three years after completion of the study. After this, all electronic materials (both original and backup copies) will be destroyed by deleting them from the storage area. All hard copies (both original and backup copies) will be shredded.

**Summary**

This chapter discussed the methodology used in this research study, describing the overall research design plan used to investigate the research question. It covered the theoretical bases of the development of the research design, followed by the description of the research design model, including an overview of the three phases of the study: the *Exploration, Development* and *Evaluation Phases*. This was followed by a description of the target population and setting, justifying the selection of the units and unit teams. The different data collection sources and methods were described for each of the study phases, and the data analysis methods were presented. The data collection methods are further discussed in Chapters 5, 6 and 7, while the summative research findings are discussed and analysed in Chapters 8 and 9.
CHAPTER FIVE
EXPLORATION PHASE
DESCRIPTION AND FINDINGS

Introduction

This chapter describes and presents the findings from the first phase of the study, the Exploration Phase. This phase of the study involved the first two stages of the HPT model: Performance Analysis and Cause Analysis. Its findings gave rise to a set of current best performance guidelines with recommendations for areas of improvement and suggested strategies to improve and enhance the actual performance of the task assessment process. The findings from the Exploration Phase guided and informed the next phase of the study, the Development Phase. The Development Phase is described and the findings discussed in the following two chapters (Chapters 6 and 7).

The Exploration Phase had limitations in scope and in time. This phase was limited to being explorative, and it was limited in time to twelve weeks. The reasons for these limitations were that this phase was designed to give only a brief snapshot or description of the desired and current work practices and environment in which the task assessment under investigation occurred. The purpose of this phase was to identify possible intervention strategies built upon the literature review where performance gains could be expected at the workplace level. Only qualitative findings were possible, as no quantitative performance metrics were available on which to measure the actual or current performance of the tutors or coordinator as they carried out the task assessment process. While performance metrics would have been desirable, such development was outside and beyond the scope of the study. This lack of performance metrics, however, did highlight the need for research in this area.

The Exploration Phase involved the analysis of the task assessment process, and consisted of the Performance Analysis and Cause Analysis (stages 1 and 2 of the HPT model). The performance analysis was carried out to identify the what, that is, the actual or typical performance (current state) of the job or process, and the difference between this state and the desired (ideal) or exemplary performance requirements of the job. This difference between these two performances, the actual and the desired, is called the gap. Thus, a gap occurs when the two performances do not align, or there is a performance
problem. The cause or diagnostic analysis was carried out to determine the why of the performance gap: what specific factors and areas caused the deficiencies, and what changes or interventions could be implemented to reduce the performance gap. The focus of this phase of the study was on the identification of possible intervention areas and factors where the coordinator and tutors had some control over their actual performance, that is, what, how and where they did their work.

The findings from the performance and cause analysis of the task assessment process have been grouped under the themes based on the conceptual framework of the task assessment process, developed in Chapter 3 (see Figure 3.3). The themes were: the marking key, marking activities, skills and knowledge, and management. These components of the task assessment process were identified as those amenable to change or intervention within the scope of the study that would help to eliminate the causes of the identified performance problems or gaps. These interventions to the task assessment process were used to inform the initial development cycle of the EPSS, discussed in Chapter 6 and helped to develop a baseline of possible features that could be incorporated into the EPSS.

**Performance analysis**

The data sources for the performance analysis came from Edith Cowan University’s (ECU) policy and procedure documents on assessment, research literature and structured task observation. These data collection methods were described in the previous chapter (Chapter 4). The data gathered during the document and research literature investigation, and structured task observations, were collated, grouped and analysed by looking for themes, trends and for illustrative descriptions. These procedures and analysis relate to first stage of the HPT model as discussed in Chapter 3 (see Figure 3.1). The use of multiple sources of data improved the validity of the findings and recommendations. The combined findings obtained from these sources of data enabled the researcher to describe the desired and actual performance and to identify the performance gap.

**Desired performance**

The desired, ideal or exemplary performance of the task assessment process was determined from two sources: ECU’s policy and procedure documents on assessment, and research literature. The desired performance levels set by the institution were very
broad, and difficult to quantify, measure, and describe in detail. This placed a great burden on the coordinator and tutors when they tried to gauge their performance against the desired institutional performance. Due to the significant changes that have occurred, and are still occurring in tertiary education, only the most recent ECU’s documents and research literature were considered appropriate for analysis.

Document analysis

The purposes or goals of assessment can be broadly classified into three groups: to provide certification (assessment of learning); to improve learning (assessment for and as learning); and to contribute to quality assurance of the teaching program (see Chapter 2, Section on Assessment). The method and selection of assessment methods, the load and frequency of assessment, and the constructive alignment of the assessment task to the unit objectives were outside the scope of this study. Although these are fundamental to assessment, they tend to focus on the assessment task itself rather than the process. The task assessment process is just as significant and, if not well founded and implemented, can place these other aspects of assessment in jeopardy. In other words, for the performance of the assessment process as a whole to be exemplary, all aspects of the process need to be exemplary including that of task assessment.

The investigation of the university policy and procedure documents, and research literature, was limited to the area of task assessment where professional judgment was involved, and covered topics such as marking key design, marking activities, moderation and management. This document investigation allowed the researcher to form a general view of the current best practices involved in the performance of the task assessment process. While much research has been published on the different types or methods of assessment (see Chapter 2, Section on Assessment), research on the actual practices involved in task assessment has been limited. Areas such as the marking key design, sub-mark allocation, workflow and workload, stress and other issues are rarely mentioned in the literature on assessment. What has been published in these areas has often been generic in nature and theoretical, with limited exploration of the practical implications.

ECU’s document on the policy and procedures of assessment (Learning and Development Services, 2005) was the primary source of data, as this was the institution involved in the study. However these policies and procedures were similar to other Australian and overseas universities (University of Canberra, 2003; University of
Queensland, 2005). The policy and procedures within these assessment documents stated the (desired) responsibilities of the stakeholders involved in assessment, and the principles of student assessment. These documents covered the whole area of assessment and all stakeholders involved in the assessment process, however, this study was only interested in the specific parts of the document that related to task assessment, and more specifically, those related to the learners, tutors, lecturer, and coordinator of a unit.

ECU’s document listed the specific responsibilities of these stakeholders. These responsibilities could be seen as forming the bases of the desired or exemplary performance of the stakeholders. However, the document did not provide any operational definitions or metrics for these responsibilities. The document began with the description of what the assessment of learners should be, that is, valid, educative, explicit, fair and comprehensive, and continued listing the responsibilities of each of the stakeholders.

The responsibilities of the unit coordinator, lecturer and tutor are listed in Tables 5.1 and 5.2. For the benefit of the reader and to facilitate further discussion, the researcher abridged this policy and procedure document, but kept the document headings and numbering system when listing the responsibilities of the relevant stakeholders.

The responsibilities of the unit coordinator in regards to the task assessment process are listed in Table 5.1. They include providing opportunities for student involvement in determining the ways in which they will be assessed, putting in place moderation mechanisms to ensure consistent marking, maintaining confidentiality of student information, and reporting on assessment results.

Table 5.1
Responsibilities of the unit coordinator according to ECU’s policy on assessment

- 6.2.9 to provide, where appropriate and possible, opportunities for students to participate in identifying their learning needs, planning their learning experiences and determining the ways in which they will be assessed; …
- 6.2.12 to put in moderation mechanisms to ensure consistent marking of all assessment tasks;
- 6.2.13 to maintain the confidentiality of personal student information including assessment results, except for legitimate University purposes;
- 6.2.14 to advise the Head of School, Program Director or Course Coordinator as appropriate, of marks and grades by the due date and to attend meetings of the School Board of Examiners;
The unit lecturers and tutors have direct responsibility for implementing the task assessment policies and procedures. These policies and procedures are listed in Table 5.2. The list, although general and limited in nature, does provide some guidelines and guidance to the desired performance of the unit lecturers and tutors as they undertake the task assessment process.

Table 5.2
*Responsibilities of the unit lecturers and tutors according to ECU’s policy on assessment*

| 6.2.15 | to assess students’ work fairly, objectively, consistently and in a timely manner and to provide adequate feedback about performance; |
| 6.2.16 | to provide timely feedback on assessments during the teaching session. To ensure for the regular semester that turnaround time for assessments is three weeks or less. Any increase in turnaround time for assessments requires the approval of the Head of School or Program Director and students must be advised of the increased turnaround period in advance; … |
| 6.2.19 | to make reasonable accommodation within University policy for students with a disability (see ECU, 2000, Disability Policy); … |
| 6.2.21 | to maintain the confidentiality of personal student information including assessment results, except for legitimate University purposes; |

The policy document also contains a section on the determination of marks and grades. This section covers the information that students must be given, the distribution of grades, and how the marks are to be presented to the University and to students. These procedures are listed in Table 5.3.

These ECU institutional policy and procedure guidelines on assessment outlined in Tables 5.1, 5.2 and 5.3 were typical of all similar guidelines on assessment reviewed by the researcher. Many institutions, including ECU, have recently published explanatory and informational documents, and offer training, especially to new staff and tutors, to improve their performance and literacy of the assessment process. A good example of this is *The Pathways to Good Practice: A guide to flexible teaching for quality learning* (Teaching and Learning Centre, 2006), an initiative of Southern Cross University’s Teaching and Learning Centre. This guide highlights a number of desired performance goals essential to quality teaching and learning, one being the ability to communicate the assessment process clearly (transparently), explicitly and concisely to both learners and assessors.
Table 5.3
Determination of marks and grades according to ECU’s policy on assessment

6.4.2
b. Students must be informed of their numerical mark, grade or pass/fail result for every component of assessment in the unit except for any final examination or test.

c. Unless otherwise approved by the Head of School, the final assessment mark for each student in the unit shall be determined on a scale of 0 - 100 per cent by the methods set out in the unit outline.

d. Examiners should ensure that marks and grades are awarded appropriately. Where grades are allocated according to a set of pre-determined standards, students must be provided with the standards prior to completing the assessment. In other situations, the recommended distribution of grades is:

* Undergraduate courses. Approximately 35% distinction passes (High Distinction and Distinction) of which not more than half should be at the upper level.

* Graduate courses. Approximately 45% of distinction passes (High Distinction and Distinction) of which not more than half should be at the upper level.

Variation around the grades will be a normal expectation. The degree of variation will be a function of the nature of the unit, the number of students and the calibre of the students.

6.4.3 Reviewing assessment marks and grades for assignments, tests and final examinations.

Summary

Based on the findings of these institutional policy and procedure documents, and the results of the literature review on assessment, the desired performance goals of the task assessment process were developed. These desired performance goals were grouped using the themes that developed from the assessment process framework discussed in Chapter 3 (Figure 3.3): the marking key, the marking activities, and the skills and knowledge of the assessor.

The desired performance goals of the marking key derived from this phase of the study indicated that the marking key should be valid, consistent, fair and reliable. The marking key should also be transparent and promote learning; and when appropriate students should be involved in formulating some or all of the assessment criteria used in the marking key.

A number of desired performances goals were identified within the marking activity theme. These were: the assessment of students’ work should be fair, objective and done in a timely manner; feedback should provide information about performance; and the assessment results are to remain confidential.

Finally, in the area of skills and knowledge of the assessor a number of desired performance goals were identified. Assessors need to be informed of the unit, course
and institutional procedures, practices and regulations associated with assessment; and these need to be revised regularly.

**Actual performance**

The actual performance or practice of the task assessment process at the workplace level was derived from two sources of data: the literature review on assessment, and the structured task observation of tutors’ during the task assessment process.

The researcher found a critical lack of research literature that investigated the actual practice and performance of the task assessment, although the research literature commented on the increase in stress, dissatisfaction, and workload associated with assessment for both students and staff. This lack of research into the practice of the task assessment process involved such areas as workflow, time spent marking and on the recording of feedback, the use of technology, and the management process. This paucity of research is especially surprising when one considers the significance of this activity to teaching and learning, the amount of time and resources allocated, and the amount of change witnessed by tertiary education institutions during the last two decades.

**Structured task observation**

In the structured task observation, the researcher observed the workflow practices of tutors selected from ECU’s School of Education as they carried out the marking process over twelve weeks. The structured task observation was concerned with how tutors currently performed their marking activities, and it occurred at university during normal working hours. The structured task observation consisted of both unplanned and planned observations: the unplanned observations involved opportunist observations of tutors marking, while the planned observations involved five selected tutors from different units of study within the School of Education being observed in the process of marking students’ work over twelve weeks. The results of these observations are discussed under the following headings: *Marking key*, and *Moderation and Management*.

**Marking key**

None of the marking keys observed during the structured task observation period met the desired design features of a marking key as described earlier, as they lacked explicit performance criteria and standards or levels of achievement. They tended to
have marks attached to simple word descriptors such as *pass*, *good* or *very good*, or had grades attached (see Figure 5.1). Even when informed by the tutors that moderation had occurred, the researcher observed tutors still had difficulty with the allocation of marks. Often at the beginning of the marking process, the tutor, after marking a number of papers, would have to readjust these marks as their initial mental model of what constituted the different qualities and values of work changed. Also observed was the issue of time between marking sessions: the further apart the sessions, the more time the tutors needed to remember and reformulate the mental model of the allocation of marks.

![Figure 5.1 Example of a criterion and word levels.](image)

The use of numbers or percentages for either sub-marks or totals has not been recommended in the literature when professional judgement is involved (Bridges, 1999). This is because the distribution and discrimination of marks tend to be distorted by this process. Bridges et al. (1999) found that when professional judgement is involved in assessment and a number system is used, the range of marks is narrowed and fewer very low or high marks are given. This is because the actual marks given tend to be meaningless and difficult to justify, for example, the difference between 54 and 55 or between 90 and 91 is not clear when a subjective judgement is being made.

**Moderation and management**

The researcher observed during the structured task observation of tutors carrying out the task assessment that when moderation did occur before marking commenced, it was rudimentary. The tutors felt that the *number* or *word-line* method used in the marking key design was a contributing factor to this difficulty in moderation. Usually the tutors marked in isolation and only contacted other staff in the unit, usually the coordinator, if the student was failing. They sent in their marks on request and the coordinator carried out any adjustments on the spread of grades and marks. All tutors observed found the *number* or *word-line* method of marking difficult to implement as they could only develop a sense of an appropriate range, for examples, what 6 out of 10 means, after marking a substantial number of assignments. This meant that they often had to go back and readjust the marks after developing a sense of appropriate allocation of marks.
The management of the marking process was highly variable but did consistently show a lack of appropriate use of technology. The researcher often found it difficult at times not to suggest appropriate procedures and uses. For example, most tutors observed when using a spreadsheet to record marks did not take full advantage of the features available within the spreadsheet that a program like Microsoft Excel offered. Tutors were observed handwriting total marks against student names on a hard copy of a spreadsheet and forwarding this to the coordinator.

The areas of moderation and management are also under-researched in terms of how to carry them out and when to use different methods, forms or procedures. The research literature on assessment seems to often assume that the reader knows how to carry out moderation and marking, and how to manage the task assessment process well, and that this is common knowledge and does not need to be researched. Yorke (1998) found that very little literature had been published on assessment management in contrast to methods of assessment. This seems surprising in light of the growing trend towards larger class sizes, and students being considered as consumers and all that this entails in terms of consumer rights and litigation. Meanwhile, Knight (2002b) extends this need for research into management to cover not just the unit of study but the study programme as a whole.

### Performance gap

The lack of quantitative performance metrics for the task assessment process meant that the researcher could only discuss the results of this exploration of the performance gap in qualitative terms. The results of the performance analysis identified significant gaps in the performance of the task assessment process at the workplace. These gaps were grouped around key areas derived from the assessment process conceptual framework discussed in Chapter 3 (Figure 3.3): 1) assessor’s skills and knowledge of assessment; 2) marking key design; 3) marking activities, including moderation feedback and reporting; and 4) management. The gaps identified are grouped and listed in Table 5.4. The magnitude of the performance gap is indicated by a number range of 0 to 5, 0 meaning no gap and 5 meaning very significant gap as judged by the researcher. This range, although based on limited data, did allow the researcher to allocate a degree of importance to the gap.
Table 5.4
*Performance gap range on key desired performance goals of the task assessment process*

<table>
<thead>
<tr>
<th>Desired Performance Goals</th>
<th>Performance Gap Range (0 to 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assessor knowledge and skill:</strong></td>
<td></td>
</tr>
<tr>
<td>Need to be trained in assessment procedures</td>
<td>4</td>
</tr>
<tr>
<td>Part-time assessors need to know about assessment practices and regulations.</td>
<td>4</td>
</tr>
<tr>
<td>Assessment practices need to be revisited regularly</td>
<td>4</td>
</tr>
<tr>
<td><strong>The marking key design:</strong></td>
<td></td>
</tr>
<tr>
<td>Should be seen to be valid</td>
<td>4</td>
</tr>
<tr>
<td>Should be consistent, fair and reliable</td>
<td>4</td>
</tr>
<tr>
<td>Should be transparent</td>
<td>4</td>
</tr>
<tr>
<td>Should promote learning, not just measure what has been learnt</td>
<td>4</td>
</tr>
<tr>
<td>Students should be involved in formulating at least some of the assessment criteria</td>
<td>5</td>
</tr>
<tr>
<td>To make reasonable accommodation within University policy for students with a disability (6.2.19)</td>
<td>5</td>
</tr>
<tr>
<td><strong>Marking Activities:</strong></td>
<td></td>
</tr>
<tr>
<td><em>Moderation</em></td>
<td></td>
</tr>
<tr>
<td>Mechanisms to ensure consistent marking of all assessment tasks</td>
<td>4</td>
</tr>
<tr>
<td><em>Marking</em></td>
<td></td>
</tr>
<tr>
<td>To assess students' work fairly, objectively, consistently and in a timely manner and to provide adequate feedback about performance (6.2.15)</td>
<td>4</td>
</tr>
<tr>
<td><em>Feedback/reporting</em></td>
<td></td>
</tr>
<tr>
<td>To provide timely feedback on assessments during the teaching session.</td>
<td>3</td>
</tr>
<tr>
<td>To ensure for the regular semester that turnaround time for assessments is three weeks or less (6.2.16)</td>
<td>3</td>
</tr>
<tr>
<td>To maintain the confidentiality of personal student information including assessment results, except for legitimate University purposes (6.2.21)</td>
<td>4</td>
</tr>
<tr>
<td><strong>Management:</strong></td>
<td></td>
</tr>
<tr>
<td>Needs to be implemented at all stages and integrated between stages</td>
<td>5</td>
</tr>
</tbody>
</table>

A number of other issues raised in the performance analysis were outside the scope and influence of this study, for example, the ICT infrastructure, training and resources, and the quality and type of assessment method to be used.

**Cause analysis**

The cause analysis identified possible causes for the performance gaps or suggested changes or interventions in the task assessment process that might reduce the identified performance gaps described in the previous section. The proposed intervention strategies emerged from the gap analysis and from the expert interviews discussed below.

In line with the assessment process conceptual framework (Figure 3.3), these proposed strategies, changes or interventions are grouped and discussed under the following four headings: 1) skills and knowledge, 2) marking key design, 3) marking activities, and 4) management. These initial areas of possible intervention where the
ones considered to be the most significant and the most likely to improve the performance of the task assessment process, and they formed the bases for the start of the first development cycle of the next phase, the Development Phase.

**Best practice interviews**

Six experts in the field of assessment and moderation were selected by the supervisor and reviewers of the research proposal. These experts were interviewed individually using a semi-structured interview protocol. The focus of the interview was on current best practice and methods used in the task assessment process. Each interview took approximately one hour to complete and was recorded on audiotape for later analysis, and notes were taken by the researcher. The data collected from the interviews were used to inform the cause analysis. The findings were analysed and interpreted by looking for themes and trends.

Three themes and trends emerged from the interview data. They were: 1) tutor training, 2) marking key, and 3) marking moderation. The experts felt that these needed to be incorporated into the task assessment process for desired performance to be achieved.

The first theme that emerged from the interview discussions was tutor training and resources. The experts identified the need for not only general training in task assessment procedures, but also the need for specific training. This specific training related to the understanding and interpretation of the marking key prior to its use. This training needed to be efficient and effective, as many of tutors were now part-time (sessional) and to find the time in which to carry out the training was difficult, as was finding the funds.

They also identified a need for tutors to improve their generic assessment knowledge and skill that related to the discipline, and suggested a generic marking scheme within the main marking key that would cover issues that are held by the discipline as general standards (e.g., referencing, grammar, form, layout, etc). Because these are the concepts students are required to do routinely within the discipline, this needed to be developed at the unit and program level. The idea behind this generic marking scheme is that the more these generic discipline skills are clarified in the marking key, the fewer marker’s judgments will have to be made during the marking. These general principles would apply irrespective of the particular assignment, and the seriousness of the infringements would increase per semester. The more these general
principles were identified and understood by the tutors and students, the more transparent the actual marking key would be. Another variation to the marking scheme that was suggested was a global sub-mark and penalty sub-marks. These types of marking schemes could be combined and incorporated into the concept of the generic marking scheme.

The second theme was that of the marking key. The best practice interviewees discussed the reasons why the marking key may not produce the desired intra-marker comparability. They suggested that it could be due to variations in the adequacy of the marking key itself and/or the adequacy of the knowledge and experience of the assessors (see previous theme on training). This knowledge and experience refers to the unit and the programme, and not to just specific knowledge and experience of the content (Lederman & Gess-Newsome, 1999). If it were assumed that the marking key is adequate and the marking shows low intra-marker comparability, this would indicate the need for education and training of the markers. However, where all markers’ knowledge and experience is high, this could produce high intra-marker comparability even though the marking key is inadequate. Thus, the more inexperienced the marker, the more detailed the marking scheme needs to be to achieve high intra-marker comparability.

Also discussed during the interviews were the different marking schemes, such as soft, preferential, cumulative and sequential marking. The difference between soft marking and preferential marking was raised, and it seemed that experienced lecturers or coordinators sometimes adopted a strategy where marking stringency increased from first to last assignment, and from first to last year of study. This was labelled as preferential marking to distinguish this type of marking from soft marking, where the mark is inflated for no educational reasons. A basic or cumulative marking scheme was defined as a list of points with marks that are simply added up, while in a sequential marking scheme, the marks build up and the student must gain marks in the previous point before being able to proceeding to the next point.

The third and final theme was moderation and was not limited to the marking activity. The involvement of tutors in the development and/or refinement of the marking key was suggested by the experts during the interview as a means of improving the quality of the marking key. The moderation process seemed to touch on all aspects of the interview discussion from training to marking. All interviewees considered moderation needed to be part of the tutor and coordinator assessment training. They also
suggested that moderation needed to occur at all stages or activities of the task assessment process, and highlighted the importance of group discussion. They suggested the use of ICT to facilitate collaboration and discussion, and pointed out how email and discussion boards had reduced the need for face-to-face meetings. This fitted in with their concern about recent changes to university culture and how this was impacting on the assessment process. They believed that the on-campus culture was changing, as both student numbers and sessional staff numbers increased, and many full-time students worked part-time or even full-time. There was thus less student-staff and staff-tutor contact in this new campus culture. They described how these changes were affecting both assessment and the moderation process.

Interviewees described how in the past, there were more opportunities (e.g., over coffee in the staff room) for co-ordinator and tutors to participate in an on-going clarification process, which involved assessment, marking key and moderation issues. However, currently there was less opportunity for these informal discussions to occur, as sessional staff numbers were increasing and meeting times were thus more difficult to organise, and these meetings were either under funded or not paid for at all. This loss of informal interaction was considered significant by the interviewees to the overall quality of the assessment process. They felt that new ways needed to be found to replace these lost informal interactions.

### Suggested intervention strategies

A number of possible broad intervention strategies were developed out of the analysis of the best practice interviews that could contribute to improved performance of the task assessment process at the workplace. These possible intervention strategies are grouped under the main areas of concern developed from the conceptual framework, and are listed below:

1) Marking activities:
   - Development of a common understanding of the assessment standards as part of the moderation process;
   - Where appropriate, development of two work samples per grade for use in the moderation process;
   - Improvement of the quality of information contained in feedback and reporting to both the student and staff through a process of moderation and review; and
   - Application of an EPSS to reduce and remove the administrative and clerical activities associated with the marking activities.
2) Marking key design:
   • Development of a moderation process for the marking key;
   • Development of a generic marking key to cover a number of units/years/courses;
   • Involvement of the markers (tutors) in the development and/or refinement of the marking key; and
   • Incorporation of instructional rubric features into the marking key to improve the transparency and educative function.

3) Skills and knowledge:
   • Provision of training for tutors in use of the marking key;
   • Provision of support to tutors in adjusting their standards to that of the marking key; and
   • Improvement of tutors’ access to the tacit knowledge on assessment held by the coordinator, through the innovative application of technology.

4) Management:
   • Improvement of integration at both the micro and macro levels of management of the assessment process;
   • Application of technology, e.g., email and discussion boards, in order to optimise the effectiveness of meetings and to facilitate collaboration with tutors; and
   • Development of moderation processes, both quality assurance and control, across the whole task assessment process.

The first group of intervention strategies (marking activities) tend to be located within the organisational systems and cognitive support area. These involved marking, feedback and reporting interventions strategies to improve performance. The findings showed that these areas were under-researched and thus had the potential of significant performance gains especially when combined and integrated within an EPSS. The use of an EPSS within the task assessment process through the incorporation of the marking key could improve the reliability, management and workflow, and eliminate or reduce the burden of the clerical and administration work involved. The structured task observation highlighted many areas where ICT and EPSS applications could be employed to improve and enhance performance of the task assessment process. However, the observed low level of ICT skills of the tutors (both part- and full-time) and of most staff within the School of Education at Edith Cowan University were also taken into consideration in the next phase, as were the difficulties of interfacing with the Edith Cowan University IT infrastructure experienced by the tutors.
The second group of intervention strategies (marking key design) are located within the theme of cognitive support and job aids. They focused on the improvement of the marking key design and the incorporation of the EPSS into the task assessment process. The use of instructional rubrics or criteria-based marking guides was suggested as the basis for the creation of a more transparent, reliable, valid, fair, educative and informative marking key for both tutors and students. The majority of current marking keys used by the tutors and observed by the researcher in this phase of the study were not transparent, reliable, valid, fair, nor were they educative or informative. They provided little feedback or guidance to either the student or tutor. The researcher found that the criteria used in the current marking keys did not describe any real standards. This meant that both tutors and learners found it difficult to tell what standard was expected or would be considered adequate or appropriate for gaining a particular mark or grade. These marking keys without any clear standards introduce a high degree of unreliability into the marking process. They also prevent learners from knowing what standards are expected and thus internalising them, and from developing the ability to monitor their progress prior to submitting the task. An educative and instructional rubric, on the other hand, provides the tutors and learners with a neutral and less emotive area in which to discuss and communicate the desired outcomes of the assessment task (Andrade, 2005; Burley & Price, 2003; Montgomery, 2002; Saddler & Andrade, 2004; Stefl-Mabry, 2004; Tierney & Marielle, 2004).

The development of the marking key should involve at least the tutors in moderation and review of the marking key. At best, the development would be a collaborative process involving the learners, tutors and coordinator. The development of clear and explicit criteria and grade descriptors would greatly enhance the transparency, reliability, validity, and fairness of the task assessment process. This would allow both tutors and learners to better interpret and understand what the assessment task requires of the learner. To achieve this better understanding, the findings suggested expanding the moderation process (quality control and assurance) of the marking activity to cover the marking key design and development. This would include the marking key development and the post marking activity. These moderation activities would ensure that the desired performance outcomes were being monitored and achieved at the desired performance level. The application of an EPSS could be employed to improve and enhance the performance of these moderation activities.
The third group of intervention strategies are located within the skills and knowledge area, and although to some extent they are outside the scope of this study, they are listed as part of the Exploration Phase of the study. They do represent the bigger area of knowledge management that the EPSS would endeavour to incorporate.

The final group of possible areas of interventions were grouped under management, and they are located within the organisational systems and cognitive support areas. The researcher found that the management of the assessment process was haphazard at best or nonexistent. Thus, performance could be improved through a review of management process both at the micro and macro levels of the task assessment process, and the application of an EPSS.

**Role of ICT**

Due to the significance that ICT would play in the study, the researcher used the four *activities of use* classification system (Hackos & Redish, 1998) to group the tutors when using ICT both during the Development Phase and throughout the study. These four *activities of use* were: novice, advanced beginner, competent performer and expert. These groups are not static but change between software application and over time, with some users not advancing beyond novice or advanced beginner. The authors (Hackos & Redish, 1998) found this classification was useful both in fieldwork observation and in the design of a product.

The tutors observed by the researcher were both part-time and full-time, and used a range of ICT methods for marking, recording of marks and feedback, and recording of students’ marks. These ranged from almost no use of ICT to a moderate level of ICT usage. Most of the tutors observed by the researcher used ICT at the *novice* or *advanced beginner* level; none of the tutors observed used ICT at the *expert* level, while some could be classed as *competent performers*. The researcher also found that tutors could not, or found it very difficult to, integrate material between the applications they were using. For example, they found difficult the movement of data from a word-processed document to a spreadsheet document and vice versa.

The processes that the tutors followed and the amount of ICT used during the task assessment process were in many of the cases determined by the coordinator. This was because the coordinator usually supplied the blank marking key, either hard copy and/or an electronic copy, and required the list of marks to be returned to them in a certain manner. This returning of tutorial lists of marks was very varied and ranged from hand-
written lists, to email-attached spreadsheets. Often the coordinator did not receive any marks until the end of the semester and, thus, could not monitor the student or tutorial groups progress with any degree of precision during the semester. Furthermore, usually tutors only interacted with the coordinator if they were having difficulties in the assessment of a student.

The tutors all indicated that they were embarking of the journey of employing ICT in their work practices and the task assessment process. The researcher found that time constraints (‘had to get the job done’) meant that they had often no time to explore how ICT might help them complete the task at hand more efficiently. Their limited knowledge of the software applications being used was also an impediment to their successful application of ICT to the task assessment process. As the tutors had learnt most of their ICT skills on the job, many were unaware of the potential power of the software applications they were using and, thus, how useful these applications could be to improve their performance in the task assessment process. For example, many tutors found downloading a comma-delimited text document of their tutorial class from the Internet, containing student details, and opening and saving this to a spreadsheet, a difficult process. Even the manipulation of this content within the spreadsheet was difficult for them. Thus, the researcher often saw the tutors using printed class lists from the Internet with hand-written additions and deletions of students. This method gave them no ability to electronically edit or add information.

In conclusion, most tutors observed by the researcher were either novice or advanced beginners. The tutors’ typical use of ICT in the task assessment process involved printing out a word-processed marking sheet and manually recording names, comments, sub-marks, and the total mark. When recording final marks, the tutors’ use of ICT ranged from writing by hand on a word-processed or spreadsheet printed document and mentally adding them up or using a calculator, to recording the marks in a spreadsheet manually but having the different assignment marks added-up electronically.

**Conclusion**

The results from this Explorative Phase highlighted the fact that current research provides only limited guidance in terms of actual practical methods for achieving the desired performance of the task assessment. However, the actual performance falls well short of these desired performance levels. A number of areas and strategies to reduce
this gap were identified where change or intervention could lead to reduction of the gap in performance of the task assessment process. These areas were: 1) marking activity, 2) marking key design, 3) skills and knowledge, and 4) management.

These suggested areas and strategies were used to form the bases on which to begin the next phase of the study, the *Development Phase*. The following two Chapters (Chapters 6 and 7) describe and discuss the findings from the six development cycles.
CHAPTER SIX
DEVELOPMENT PHASE – SEMESTER 1
DESCRIPTION AND FINDINGS

Introduction

The Development Phase consisted of six collaborative iterative development cycles completed in an authentic work environment over two semesters in 2003. This chapter provides an overview of the Development Phase, describing the software platform used, the background of the team members and the units selected for this study. This chapter also provides an overall description of the structure of a typical development cycle. Finally, this chapter describes and discusses the findings of the first three development cycles that occurred in Semester 1. The next chapter (Chapter 7) describes and discusses the findings of the three development cycles that occurred in Semester 2.

Each development cycle (DC) is divided into task assessment activities, usually seven. These activities were built upon the findings from the Exploration Phase as being significant activities within the task assessment process, and they are used to help in the discussion of the findings. To further facilitate an understanding of the development and evaluation of each development cycle, the following structure is used to describe each DC: 1) an introduction that sets the scene and describes the assessment and objectives, 2) the detailed description and analysis of the team members’ role within each activity, and 3) the findings from each DC.

Development Phase overview

This section briefly covers the rationale for the selection of the software development platform, the team members involved in the research study and their educational and ICT background, and a brief description of the units selected for this study.
Development platform

The software development platform selected to develop the EPSS was FileMaker Pro, a relational database system. Relational databases allow the sharing of data between datasets of files based on a linking or key field. This means that data need only be entered once but is accessible from other databases, thus making the data more easily kept up-to-date and managed. FileMaker Pro is a sophisticated and mature relational database system that combines spreadsheet and word processing features, such as calculations, spell-checking and mail merge. It is highly scalable and used in commercial products, and is “the easiest relational database system to use” (Stars, Child, & Bernard, 2004, p. 3). Furthermore, this platform was selected because the researcher was familiar with the programme, and because FileMaker Pro databases could be used on the most common PC and Mac computer operating systems. Finally, an additional and significant reason for choosing FileMaker Pro was a cost-saving feature: the ability to create runtime copies of the EPSS application without having to install copies of FileMaker Pro on each computer.

Team members

The unit team in each semester consisted of the unit coordinator and four tutors. Table 6.1 shows the members of each unit team. They were all females, each with over 20 years experience in the teaching profession, with a number having overseas experience in less developed countries.

Table 6.1
Unit team members

<table>
<thead>
<tr>
<th>Semester 1 – EDL1101</th>
<th>Semester 2 – EDL1201</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinator</td>
<td>Coordinator</td>
</tr>
<tr>
<td>Tutor A (part-time)</td>
<td>Tutor A (part-time)</td>
</tr>
<tr>
<td>Tutor B (part-time)</td>
<td>Tutor B (part-time)</td>
</tr>
<tr>
<td>Tutor C (full-time)</td>
<td>Tutor E (full-time)</td>
</tr>
<tr>
<td>Tutor D (part-time)</td>
<td>Tutor F (full-time)</td>
</tr>
</tbody>
</table>

The summative findings resulting from Semester 1 (first three development cycles) were incorporated into the design and development of the EPSS that was used at the start of Semester 2. The team members involved in the Semester 2 unit (Coordinator, and Tutors A, B, E and F) were the ones interviewed at the end of the Development Phase. The findings of those interviews are part of the Evaluation Phase and are discussed in Chapters 8 and 9. Therefore, to help with the understanding and
interpretation of the findings, a brief description of the Semester 2 team’s educational and ICT experience is given below.

The coordinator of both units had over 25 years teaching experience across all levels of education, predominantly early childhood in disadvantaged settings, both in Australia and overseas. Her professional experience involved educational leadership roles at school (Principal), district and state levels, having lectured at university for the last three years. Her ICT experience had mainly involved on the job learning of Microsoft Office and other applications such as EndNote and Inspiration. Her ICT ability was at the novice to advanced beginner level.

Tutor A had worked for over 25 years as a primary school teacher. During the last six years up to the end of 2003, she had worked with pre-service teachers who were completing their Professional Experience. The year of the study was her first tutoring experience at university. Tutor A had no prior ICT experience and felt hesitant about using technology; this placed her below the novice level.

Tutor B had worked in the field of primary education for over 20 years. She began teaching as a general primary school teacher for the first six years and then as an education consultant for the Education Department of Western Australian (EDWA) for two years. During the 1990’s, she worked part-time as an Arts teacher. In 2000, Tutor B spent one year overseas as a primary school principal with Australian Volunteers Abroad (AVA). The year of the study was her first tutoring experience at university. Her ICT skill was at novice level and was limited to basic Microsoft applications, with limited experience and some professional development in using digital media. Her professional role as deputy principal and principal at a primary school had brought her into contact with databases, while her use of the Internet and email was limited at the time of the research study.

Tutor E had worked in the field of education for over 25 years. She had worked in secondary education for the first eight years; the next year was spent as an education consultant for EDWA; and the following 17 years involved work in the tertiary area, including sessional, contract lecturing, research and Practicum management. She had been using ICT since 1988 and was familiar the Microsoft applications and the Internet. Her ICT level of experience was judged to be advanced beginner to competent performer.

Finally, Tutor F had taught for over 30 years at all levels of education. Her professional career began with three years in primary education, followed by two years
in the South Pacific teaching high school. The next 11 years she was involved as specialist English as a Second Language (ESL) teacher with EDWA, and the following two years she worked as an educational consultant with EDWA. Tutor F was a full-time student over the next two years, and she was currently tutoring and lecturing at university. She had been using ICT since the late 1980’s and had a good command of Microsoft Office, Endnote and Internet applications. This experience gave her an ICT level of competent performer.

Although the team members were all very experienced teachers and many had professional development experience, a number of them were new to tertiary education. This broad range of teaching experience contributed greatly to this research study. Their ICT experience ranged from below novice to competent performer level. This broad range of ICT experience meant that the team members were representative of the wider teaching community, thus enhancing the validity of the results of the study.

Due to the successful rapport and collaboration that was built between the researcher and the other team members during the Semester 1 development cycles, both the unit coordinator and two of the tutors were willing and able to continue with the development of the EPSS in the second semester. Thus, of the original team members from the Semester 1 unit, three were able and willing to continue with the development of the EPSS. This team continuity resulted in a greater level of sophistication in the design and features of the EPSS than would have been possible had the Semester 2 unit started with a completely new team.

**Description of units**

The units selected for the Development Phase were EDL1101 Learning and Development I in Semester 1, and EDL1201 Learning and Development II in Semester 2. The EDL1101 unit team consisted of four tutors (one full-time and three part-time) and one coordinator, with 237 students enrolled. The EDL1201 unit team consisted of four tutors (two full-time and two part-time) and one coordinator, with 218 students enrolled. The assessment of both units consisted of three assessment items, all of which involved professional judgement. Table 6.2 shows a summary of the assessment tasks for each unit, with an indication of the marks allocated to each and an overview of the assessment task.
Table 6.2
Assessment tasks and development cycles by unit

<table>
<thead>
<tr>
<th>Semester 1 Unit EDL1101 – Learning and Development I</th>
<th>Development Cycle</th>
<th>DC1</th>
<th>DC2</th>
<th>DC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Task</td>
<td>Tutorial Oral Presentation</td>
<td>Poster Presentation</td>
<td>Essay</td>
<td></td>
</tr>
<tr>
<td>Mark</td>
<td>30%</td>
<td>30%</td>
<td>40%</td>
<td></td>
</tr>
<tr>
<td>Task Description</td>
<td>A collaborative oral presentation of 20 minutes, involving a team of three students.</td>
<td>A one-metre poster to demonstrate the students’ understanding about the way in which people learn.</td>
<td>A 1500-word essay in which the student was to assert their understanding of the process by which learning occurs, and provide supporting evidence from the reading and research regarding the theories he/she had elaborated on in their poster presentation.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Semester 2 Unit EDL1201 – Learning and Development II</th>
<th>Development Cycle</th>
<th>DC4</th>
<th>DC5</th>
<th>DC6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Task</td>
<td>Tutorial Paper</td>
<td>Tutorial Presentation</td>
<td>Written Exam</td>
<td></td>
</tr>
<tr>
<td>Mark</td>
<td>20%</td>
<td>30%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Task Description</td>
<td>A 1000-word tutorial paper on the importance of the social/emotional dimensions of learning to be group peer-assessed during the Week 5 Tutorial and then moderated by the tutor.</td>
<td>Identify, present and analyse a typical learning event the student has seen in their practicum setting. A concept map technique will be used to present their information on an A3 sheet of card or paper. The ‘event analysis’ map will be presented to a group of students in the tutorial, and handed in to the tutor at the conclusion of the presentation.</td>
<td>Development of written responses to two questions selected from a set of questions published during the semester. The first question was compulsory and was worth 30% of the total mark, while for the second question, the students had a choice of two questions and it was worth 20% of the total mark.</td>
<td></td>
</tr>
</tbody>
</table>

Each assessment task constituted one DC. A description of the activities involved in a typical DC is provided in the next section.

**Description of a typical development cycle**

The development of the EPSS and strategies for each of the six assessment items (three per semester) went through a similar DC that took between four to six weeks to complete. Each DC was divided into seven activities: Activities 1 to 6 took approximately one to two weeks to complete, while Activity 7, which involved issues and strategies associated with the management of the task assessment process, occurred throughout the DC.
The seven activities are listed and are described below with a summary of the roles of the coordinator, researcher, unit team and tutors. These roles slightly varied throughout the Development Phase, however the general workflow did not:

Activity 1  Marking key design and development
This involved the design and development of a draft version of the marking key rubric by the coordinator and researcher, based on the assessment objectives.

Activity 2  Marking key quality assurance
This involved the tutors reviewing and modifying the draft marking key first individually and later at a group moderation meeting at which the final version was decided upon.

Activity 3  Incorporation of marking key within the EPSS
This involved the researcher incorporating the final version of the marking key into the EPSS, and preparing the tool for use by the tutors for marking.

Activity 4  Pre-marking moderation
This followed the traditional moderation practice of marking a number of students’ samples by the tutors, and discussing the marks and coming to a consensus on the allocation of marks based on the marking key.

Activity 5  Marking
This involved tutors using the EPSS in the traditional marking process.

Activity 6  Post-marking moderation
This involved the moderation of the final marks (quality control), and the moderation of the tutors’ feedback comments (quality assurance).

Activity 7  Management
This involved both the management of the task assessment process and the combining of all the task assessment results to arrive at the final unit mark.

During the six DCs, the activities did not change greatly, however, what the unit team members did, that is, the processes changed as a result of the interventions and strategies incorporated into the EPSS throughout the DCs. Some of the changes and interventions were influenced by the type of assessment and how the assessment was implemented, while others were due to the iterative nature of the Development Phase and the feedback, comments and suggestions from the team. These interventions and strategies are discussed under each of the six DCs later in this chapter and in the following chapter (Chapter 7).

At times, the DC activities involved either individual tutors or the unit team working together outside the regular unit meeting times, while some activities partly took place during the unit team regular weekly meetings, which had some time allocated to the DC process when the need arose. The activities occurred primarily
sequentially, however, they did overlap due to tight time constraints and because team members worked at different paces and on different activities. This sequencing and overlapping within each DC is illustrated in Figure 6.1. Additionally, in both semesters, the researcher observed that during the end of semester the last two DCs overlapped.

![Figure 6.1 Sequence of activities within the development cycle.](image)

The development and moderation of the marking key rubric (Activities 1 and 2) was an additional feature of the EPSS design not originally considered in the initial design concept. This feature was based on two specific findings from the Exploration Phase: the need for a more detailed and transparent marking key, and the need to involve tutors in the development and/or review of the marking key. Later in the Development Phase of the study, the tutors and coordinator explored a range of marking key innovations that the EPSS made possible or facilitated (the discussion of these innovations is part of the Evaluation Phase and is covered in Chapter 9).

The refinements, additional features and strategies resulting from one DC were usually incorporated into the next development cycle. These modifications and refinements were based on: 1) feedback from the tutors and coordinator obtained at the unit team meetings, 2) individual feedback, and 3) the observation of tutors using the tool during the marking process. Some of these refinements and features were specific to particular types of assessment and could not be carried over to the next DC, whilst other refinements and features were generic in nature and were applied to the next DC.
All these refinements, improvements and additions to the EPSS, both generic and specific, tended to result in an increase in effectiveness and efficiency of the activity and the task assessment process in general. These improvements in performance were partly due to the iterative nature of the development process and to the unit team becoming more familiar with the processes involved in the activities within each DC. This resulted in an increase in the overall efficiency of the DC and gave the tutors and coordinator more time to reflect on the marking judgements and the opportunity to explore additional innovative approaches. This collaborative iterative process produced a unique version of the EPSS for each DC.

**Semester 1 development cycles**

Unit EDL1101 (Learning and Development I), a first year unit in the Kindergarten through Primary Course, was the unit selected to start the project’s Development Phase in Semester 1 2003. The focus of the unit was on the learning process and its relationship to development. The students explored their own learning as they learned about how children learn, develop and make meaning of their worlds. The unit team consisted of the unit coordinator and four tutors, three of whom were part-time (sessional), with 237 students enrolled (see Table 6.1).

During the first unit team meeting of the semester, the researcher was introduced to the tutors by the coordinator, and outlined and discussed the aim of the research study. The unit coordinator and researcher had already established a good working relationship due to the researcher’s role as IT Support Officer for the School of Education staff, and the unit coordinator had agreed to allocate time to the researcher at the start of the meeting. This time was used to give a short description of the research study, to hand out ethics clearance forms, and to describe briefly the roles and the expectations of the team members. The researcher emphasised the importance of the tutors’ collaboration in the research study and the fact that the researcher’s role was mainly one of advisor, guide and observer on ICT issues, workflow and usability issues related to the design, development and implementation of the EPSS tool.

To facilitate communication about the research, the researcher set up a website for the EDL1101 unit which contained an overview of the research study, a link to each assessment, and material related to the assessments. This website was initially set up by the researcher to be a central point of contact for the unit team members, to give them access to information about individual assignments, and to allow them to complete the
quality assurance of the rubric marking key (Activity 2) and the pre-marking moderation activity (Activity 4) via the Internet. As will be explained later in this chapter this Internet access proved to have many difficulties mainly due to access issues, and therefore a number of solutions were investigated.

The assessment of the unit consisted of three assessment items, each of which went through a DC. These three DCs are described below in separate sections. Each section begins with a description of the assignment task followed by a description of the development cycle activities, and concludes with the findings and recommendations.

**Oral presentation development cycle (DC1)**

**Description of assessment task**

The first assessment item was a Tutorial Oral Presentation and was worth 30% of the total unit mark. The sources of data for this development cycle are summarised in Appendix A. This oral presentation consisted of a collaborative task involving a team of three students. The teams were organised during the Week 1 Workshop, and in Week 3, the teams selected the week in which they would present, the first presentations beginning in Week 4 of the semester. Presentations were to take 20 minutes for each team and had to follow the guidelines for oral presentations that were included in the unit guide (the marking key used in the previous year is shown in Appendix B). The topic of the oral presentation was ‘Learning about learning’, and each student in the team was required to develop a response to a question on learning that consisted of a number of sections (See Appendix C).

**Description of activities**

This section describes in detail the Tutorial Oral Presentation assessment DC. This description is divided into the seven activities previously described. A summary of these activities is presented in Table 6.3, showing what occurred and what the unit team members were doing during each activity of the DC.
### Table 6.3

**Tutorial oral presentation development cycle (DC1)**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Contributor(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Marking key design and development</td>
<td>Researcher</td>
<td>Placed links to the criteria headings and grade descriptors on the unit website.</td>
</tr>
<tr>
<td></td>
<td>Tutors</td>
<td>Logged onto the unit website and entered their comments, judgements, thoughts, etc, under the grade descriptors for each criterion.</td>
</tr>
<tr>
<td>2: Marking key quality assurance</td>
<td>Coordinator</td>
<td>Collated and summarised the grade descriptors to produce a new set of grade descriptors based on the online feedback from the tutors.</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>This new set of criteria and indicators was reviewed and edited during a moderation meeting and the final marking key was produced. Marks for both criteria and indicators were also allocated.</td>
</tr>
<tr>
<td>3: Incorporation of marking key within the EPSS</td>
<td>Researcher</td>
<td>The criteria, indicators and marks were incorporated into the EPSS.</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>The EPSS was presented to the team at the regular meeting for a final signing off. A hard copy of the marking key was handed out and the electronic version was demonstrated.</td>
</tr>
<tr>
<td>4: Pre-marking moderation</td>
<td>Coordinator</td>
<td>Selected samples of students’ work for moderation. The oral presentation moderation could not occur before the tutors had video a number of oral presentations.</td>
</tr>
<tr>
<td></td>
<td>Researcher</td>
<td>Converted the selected sample for the Internet and PC.</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>Marked the samples on the web before the moderation meeting. At the moderation meeting, the digital movie presentations were viewed. Using print-outs of the marking key, tutors marked the video samples of work and discussed the allocation of marks, etc.</td>
</tr>
<tr>
<td>5: Marking</td>
<td>Researcher</td>
<td>The EPSS was made ready for marking by the tutor.</td>
</tr>
<tr>
<td></td>
<td>Tutors and researcher</td>
<td>As the oral presentation ran over a number of weeks, it was not possible to mark them all at once. Tutors aimed to complete the marking within three weeks of the student presenting and the researcher helped with the quality assurance, spell checking, layout and printing.</td>
</tr>
<tr>
<td>6: Post-marking moderation</td>
<td>Researcher</td>
<td>The tutors’ databases were combined to produce a unit assessment database.</td>
</tr>
<tr>
<td></td>
<td>Coordinator</td>
<td>Quality control and assurance of marks and comments occurred between tutors, tutorial groups and the unit. The spreadsheet view and student marking sheets were printed and a pdf version was produced for back-up purposes.</td>
</tr>
<tr>
<td>7: Management</td>
<td>Team</td>
<td>Occurred throughout the task assessment process.</td>
</tr>
</tbody>
</table>

Activity 1, marking key design and development, involved the collaboration of the researcher and coordinator in the development of the working copy of the marking key rubric. A number of appropriate oral presentation marking keys were obtained by conducting a literature survey, and found that the most useful and appropriate were the ones based on a rubric model. The researcher and coordinator then obtained a number of oral presentation rubric samples from both books and the Internet. Based on the objectives of the oral presentation assignment, the sample rubrics, and using the coordinator’s experience gained from the previous time the assignment had been given, a tentative working draft copy of the marking key rubric was designed and typed into a Word document. The oral presentation rubric that was used to build the working copy
was based on the *Evaluating Student Presentations* publication developed by the Information Technology Evaluation Services, NC Department of Public Instruction (McCullen, 1997).

A number of design decisions for the EPSS were also made during this activity (Activity 1) in collaboration with the coordinator and researcher. The key decisions were:

- The *student feedback* view would be limited to one A4 screen in portrait layout;
- The *tutor marking* view would be designed to fit the landscape layout of the monitor display;
- The number of indicators/descriptors would be limited to four and be based on the university grades of pass, credit, distinction and high distinction; and
- There would be a comment box under each grade descriptor for the tutors to record their comments.

These design decisions helped to guide the layout of the working copy of the marking key rubric, and determined how the information would be entered and commented on within the limited space of an A4 portrait screen.

This working copy of the marking key rubric was then entered into the EPSS database and a web interface was produced by the researcher. This web interface formed part of Activity 2, rubric moderation, and allowed the tutors to view the marking key and enter their comments and suggestions via the Internet using a login password. The coordinator then reviewed and collated these comments and suggestions online, and refined the marking key accordingly. This revised version of the rubric was then moderated, quality-assured by the unit team at the weekly unit meeting, and any agreed changes and refinements were incorporated into the rubric marking key design. In addition, during this meeting the allocation of criteria marks and descriptor marks were decided upon and moderated. By undertaking this process of rubric moderation (quality assurance), the unit team members gained a fuller understanding of the rubric and its marking scales, and took ownership of the final version of the rubric marking key before they began to use it to mark work.

Only those activities that involved tutors were placed on the web page. These activities were the quality assurance of the marking key rubric (Activity 2) and the pre- and post-marking moderation (Activities 4 and 6). As shown in Figure 6.2, this web page had links allowing the tutor to complete the three different moderation activities on-line.
EDL1101-Assignment 1

E-Marking Moderation Process

- The QA of the marking key occurs before it is used in the moderation of marks. Please read through the marking key provided on the web, making comments and suggestions in the boxes provided.
- The moderation of marks using the web marking key. A number of sample presentations will be available for you to view and record your marks on the web for comparison.
- The final QA of the marking key is done during and after the marking of assignments. Please record any comments and suggestions in the boxes provided as they arise during the marking process. These comments and suggestions will be considered when the unit is offered again.

**Links to Unit Documents**
- Unit outline
- Handbook
- Assignment 1
- Assignment 1 guidelines
- Assignment 1 presentation marks
- Grade Disccriptors
- Graduate Attributes

*Figure 6.2 Assignment 1 home page.*

The links to the different types of moderation processes, the quality assurances (Activities 2 and 6) and the quality control (Activities 4 and 6) of the marking key rubric required a password. The quality assurance web pages displayed the proposed criteria headings and descriptors, and provided space for the tutor to record their comments and suggestions about them. The first link on the web page was to the quality assurance of the working copy of the marking key rubric, and allowed the recording of both the tutors’ comments and suggestions, and that of the coordinator’s edited summary. Table 6.4 shows the tutors’ comments and coordinator’s revised texts for each of the grade descriptors for the criterion *Organisation*.

The second link on the web page is to the marking moderation page. This marking moderation page allowed the tutors to mark a number of assignment samples of students work online (Activity 4). Figure 6.3 shows the tutors’ web page view for recording comments and marks for two criteria using one student sample.
### Table 6.4

*Example of tutors’ comments and coordinator’s revised grade descriptors (DC1)*

<table>
<thead>
<tr>
<th>Organisation Criterion</th>
<th>Pass</th>
<th>Credit</th>
<th>Distinction</th>
<th>High Distinction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original Grade Descriptor</strong></td>
<td>Audience cannot understand presentation because there is no sequence of information.</td>
<td>Audience has difficulty following presentation because student jumps around.</td>
<td>Student presents information in logical sequence which audience can follow.</td>
<td>Student presents information in logical, interesting sequence which audience can follow.</td>
</tr>
</tbody>
</table>

**Tutor B Comments**
- Physically jumps around or mentally?

**Tutor D Comments**
- Because students sequence of information does not make sense

**Tutor E Comments**
- Instead of ‘student jumps around’ – substitute – ‘there is not a consistent sequence of information’

**Coordinator’s Revised Grade Descriptors (based on oral and written feedback)**
- The logical connections between ideas are difficult to understand. Little concern for timing.
- Audience has difficulty following presentation because there is not a consistent sequence of information. Evidence of logical connection between some ideas. Some attention to timing.
- Presentation has an introduction, main part and conclusion. Evidence of attention to overall timing.
- Presentation flows sequentially into a coherent whole. Shows logical thought through the introduction, main part and conclusion. Attention to the overall timing, and the component parts within the presentation is demonstrated.

The final link on the web page was to the final quality assurance and allowed the tutors to edit and comment on the marking key after having completed the marking (Activity 6). This web page recorded feedback, both suggestions and comments, about the marking key that could be used next time the assignment was offered or where the comment was generic enough to be used in other DCs. This web page was identical to the one used to record the initial comments and suggestions about the criteria and descriptors rubric moderation (Activity 2).
Activity 3 involved the incorporation of the marking key rubric developed in Activities 1 and 2 into the EPSS by the researcher. A minimalist EPSS prototype of the oral presentation marking key was developed. This was based on the findings and recommendations from the coordinator and unit team that resulted from Activities 1 and 2, and the Exploration Phase findings. This activity also involved the design of three views or layouts. These layout views are described below, and are called tutor marking view (Figure 6.4), student feedback view (Figure 6.5) and spreadsheet view (Figure 6.6). All three layouts were designed to be viewed on the computer screen, and the student feedback and the spreadsheet views were also designed to be printed.

The first two layouts or views were different views of the marking key rubric. The tutor marking view (Figure 6.4) was used by the tutors for recording their marks and comments. This view was designed to fit the landscape shape of the computer screen, thus reducing the need for scrolling with the mouse. The marks were entered by clicking on a radio button. This button was a small circular area next to each mark that became greyed-out when selected by the tutors by clicking on it with the cursor. The EPSS was designed to automatically add these criteria marks and to display the total on the screen showing the grade level achieved. If a criteria mark was changed, the total
was amended instantly. The EPPS was also designed to allow the tutor to record their comments.

![Figure 6.4 Tutor marking view – A4 landscape (DC1).](image)

The student feedback view, the copy returned to the students, was designed to be printed on one portrait size A4 sheet with the sub-marks hidden. Figure 6.5 shows a blank copy of the student feedback view, with the sub-marks hidden but the radio buttons showing. A marked student feedback view would show the greyed-out radio button below the relevant grade descriptor, and the tutor’s comments when appropriate.

The spreadsheet view displayed a list view or spreadsheet-type view of all the students showing each criteria mark, total mark and grade for each student (see Figure 6.6). The movement between these views and other functions, like sorting, finding and printing did not require any shortcuts but relied on the menu bar features of the program (FilemakerPro) used to develop the EPSS prototype. This was a deliberate early design decision as there was no time for training built into the Development Phase and the feedback from the tutors was to keep interactions with EPSS as simple as possible.
Activity 4, the pre-marking moderation, involved the coordinator selecting a number of movies, in digital format, of the students’ oral presentations. As no previous student samples of oral presentations existed for this unit, the unit team had to wait until
a number of students had completed their oral presentations. The coordinator then selected from the recorded digital video (DV) tapes a range of oral presentations that would be used in the moderation process. The oral presentation DV tapes of these students were then converted into two versions: one for viewing on the web (small file size) and the other for viewing locally on a computer (larger file size). The web version was used by the tutors to individually moderate via the Internet, while the other version was used for both individual and group moderation.

![Spreadsheet view (DC1)](image)

**Figure 6.6 Spreadsheet view (DC1).**

In Activity 5, the marking activity, groups of usually three students presented an oral tutorial. A student was assigned by the tutor to do the videoing of the presentations. This video was used for the reviewing and marking of the students’ presentation, and in the moderation process, both pre-marking and during marking. In the marking process, all tutors used hard copies of the marking key rubric to record their notes and comments during the presentation. The tutors then used the videos to review the oral presentations while they were finalising and recording the presentation marks in the EPSS. These were finally printed and returned to the students within three weeks of presenting.

Activity 6 consisted of the final post-marking moderation process of reviewing all the assessment marks, and included quality control and assurance issues. However, in this DC, this activity did not follow the typical pattern and quality control was only limited. This was because the student oral presentations occurred over a number of weeks and thus could not be marked at the same time and be returned *en masse* within the Edith Cowan University regulation of two weeks return of assignments.

Quality control involved the moderation of the marks and comments. In this DC, moderation could only be carried out to a limited extent as the coordinator could not review the range of marks and the number within each grade, within the unit and tutorial groups before the marked assignment were returned to the students. The feedback comments were also reviewed for quality, consistency, spelling, grammatical errors, and for evaluation purposes. Quality assurance involved reviewing the marking key itself that was returned to the student. This quality assurance process involved the recording of tutors’ and students’ comments and suggestions for improving the task.
assessment process. This was part of the iterative nature of the development process and would aid the next DC.

**Discussions of findings from DC1**

In this initial development cycle (DC1) involving the *Tutorial Oral Presentation* assessment, many secondary and peripheral issues needed to be dealt with before the unit team could focus on the core issues of design, development and implementation of the EPSS. The peripheral issues were in part due to the fact many of the tutors had not met each other before and this was the first time they had tutored in this unit and/or worked at a university. This meant that most unit team members had to familiarise themselves with university procedures and administration. For these reasons, and due to the fact that the tutors were still developing an understanding of what the study involved, a limited number of interventions, recommendations and innovations were proposed by them, either as team or as individual tutors during this first DC.

The outcomes and findings are discussed in the order in which they occurred throughout the six activities. As this was the first development cycle (DC1), several issues were encountered that were not central to the main research questions under investigation. These secondary issues related mainly to Internet access that prevented tutor access to the EPSS for the quality assurance of the design and content of the rubric marking key (Activity 1 and 2) and the pre-marking moderation (Activity 4). Other issues raised by the tutors related to the use of different platforms (PC or Mac), and the use of different browsers and browser versions. These difficulties prevented access to the unit website, thus preventing tutors from recording comments and suggestions on the indicators of the rubric marking key. Initially, the sessional tutors encountered problems accessing the Internet both from home and at University. This was due mainly to a lack of familiarity with the technology (e.g., two of the tutors, both sessional, had only recently received laptops and were still familiarising themselves with their operation). A contributing factor was that they had not previously been employed at the University and had not yet received their official login and password required for accessing the network, computers and email. Another sessional tutor, although a continuing employee of the University, did not have a laptop and was extremely hesitant about using any computer technology, and was unable to log in to the research study website from home.
Even though Activity 1 built upon the prior research carried out by the researcher into rubric design, this activity presented challenges due to the constraints of the completion time (around a week) and the difficulty of finding suitable meeting times with the coordinator and researcher. The coordinator and researcher agreed that both the process and the resulting marking key were a success. However, the amount of time that was spent would need to be reduced for the activity to be undertaken under normal working conditions. A significant amount of time was spent determining the best application and layout to be used to develop the working copy of the marking key rubric. Microsoft Excel and Word were trialed, and due to the coordinator being more familiar with the word processor features, a Word document was used to hold the working copy.

The Internet access issue had a major negative impact on Activity 2, the marking key quality assurance, as all sessional staff had problems accessing the unit website. To overcome this access problem, the tutors who could not access the Internet were supplied with a hard copy of the working version of the rubric. Their comments and suggestions were recorded via hard copy before the quality assurance meeting. Photocopies of these and printouts of the web page containing the other tutors’ comments and feedback were handed out and used in the process of quality assurance of the marking key during a normal weekly unit meeting.

As the available time was limited to only about a week for the researcher to solve these Internet issues and to complete Activity 2, some of the access issues could not be fully resolved in this first DC. However, the team felt that the Internet access for Activity 2, quality assurance of the marking key, should be tested again in the next development cycle (DC2). Despite the Internet access problems experienced, the web recording process worked effectively, and a number of tutors were able to log on and complete the quality assurance process. This allowed the coordinator to collate and edit the criteria and descriptors online (this process related to one criterion is illustrated in Table 6.4). This version of the marking key rubric was the one used as the working copy at the quality assurance meeting (Activity 2). A copy was handed out at the meeting to all the tutors, and comments and suggestions were recorded by the coordinator on a master copy. At the end of the meeting the coordinator edited the draft rubric marking based on the meetings comments and suggestions, and emailed this to the tutors as a final check; this revised version of the marking key became the official copy. In addition, at the meeting, the criteria marks were allocated and the descriptor
(grade) marks were discussed and decided. During the discussion about criterion mark and sub-mark allocation, a number of issues were raised that continued throughout the Development Phase. These issues were: how many marks to allocate to each criterion and to each grade descriptor, and whether the students should see these criteria and descriptor marks.

This final marking key rubric was incorporated into the EPSS (Activity 3). The layout was based on the student feedback view; the other two views (spreadsheet and tutor marking views) were developed by the researcher at the beginning of Activity 3 during the first DC.

Activity 4, the pre-marking moderation process, involved tutors moderating online a number of student oral presentations by recording their marks (see Figure 6.3) and viewing a list of other tutors’ marks for the same presentation. However, similar problems and issues were encountered in accessing the Internet as were for Activity 2. For example, tutors were unable to view the videos due to not having the correct QuickTime plug-in and/or a slow modem speed, or not knowing how to download and install the plug-in. In addition, there was very limited turn-around-time available in which moderation could occur. This prevented some of these technical problems from being resolved. The design and development of the online view of the marking key rubric required additional time and resources on the part of the researcher to both display the rubric marking key and to convert the recorded video tapes of the students’ presentations to digital format for viewing online.

The unit team, during discussions about this activity, suggested that the design of the student feedback view be modified. The suggestion was that the partial marks be hidden or removed from this view and only the solid or filled-in radio button be shown to indicate the grade that the student achieved in that criteria, the only mark to be shown being the total mark for the assignment. The hiding of the marks required some thought by the researcher and the solution was to set the text colour of the partial marks on the student feedback view to white. When printed and viewed in the student view the marks were hidden. Only the total mark and grade were visible.

Suggestions from the unit team for the online marking key moderation process were to remove the horizontal scrolling, and to consider doing the moderation in hard copy mode only. However, the unit team felt that although most team members had problems with accessing the Internet during this activity they would give this process of online moderation another try in the following development cycle (DC2). Another
suggestion about the marking key design was to reduce the number of comment boxes from one per grade descriptor to just one per criterion. This was implemented in DC2.

During Activity 4, pre-marking moderation, quality control was only partly successful with the online moderation process unsuccessful due to access and time issues. However, the viewing of the digital movies in a group meeting using a data show projector did prove to be successful. The tutors marked the video presentations using hard copies of the marking key rubric and then discussed the marks using a normal moderation process. Based on the knowledge gained from the moderation process, the tutors began the marking of the presentations.

Activity 5, the marking activity, was a lengthy process as only two presentations (involving a group of three students per presentation) could occur during each weekly tutorial. As a result, the tutors were not able to mark all the presentations at one time due to the University requirements that assignments be returned to the students within two weeks of them being handed in.

Out of the five team members, only three initially showed interest in using the EPSS to record the students’ marks and comments electronically, as they reviewed and marked using the recorded videos of student oral presentations. However, as the other two tutors observed this process being carried out by the three tutors and in conversations with these tutors, one of them decided to try using the EPSS to record marks as they reviewed the videos of the students’ oral presentations using a recently acquired laptop. The tutor who decided not to use the EPSS was a very novice user of ICT and did not have easy access to a computer. This tutor commented later, after using the EPSS for marking the second assignment, that “I wish I had [used it in the first assignment]”.

All the tutors who used the EPSS during this drawn-out marking period reported to the researcher both during the weekly meetings and individually that they found the marking key rubric and the EPSS method of recording marks reduced the boring and busy work of adding up and recording. In addition, they indicated that the use of the videos of the students’ oral presentations helped them to be more consistent with their marking. The tutors found that the rubric helped improve the consistency of both the intra- and inter-tutor marking. Additionally, the EPSS allowed for the efficient fine-tuning of the sub-marks and comments without the need to reprint the rubric. The EPSS also allowed the tutors to edit and spell-check their comments. The ease and efficiency with which this could be done was a feature that the tutors all commented upon
favourably. The tutors did however, require minor help in using the spell-checker and in
printing the student reports, as this was the first time they had used a database and they
were unfamiliar with the functions available within such an application.

As none of the tutors had used a database before to record data, they found the
auto-save and the different views of the same data difficult to understand and at times
somewhat confusing. Unfortunately, there was no time for training, except in the very
limited meeting times and in short individual meetings with the researcher. However,
the researcher found that this did not constitute a major issue, as the tutors and
coordinator were always willing to persevere and to seek help and guidance when
needed. This positive attitude from the team members contributed greatly to the success
of the research study.

Activity 6, the post-marking moderation, presented specific challenges. As the
marks were returned to the students gradually, it was impossible to conduct a post-
marking moderation of the whole unit. However, as the tutors were ready to return a
group of students’ marks, these were reviewed, discussed and moderated at a unit team
meeting. Although the post-marking moderation was restricted, the coordinator and
researcher did look at the spread of marks overall, between and within tutorial groups
and found that the spread was acceptable. The coordinator found the process of
obtaining a complete set of student results from all the tutors a much more efficient
process compared with previous methods they had used. Additionally, team members
found the following features most useful and helpful: being able to view and print
without any additional work being required; the adjustability and adaptability of the
students’ marking sheets where the sub-marks could either be hidden (student feedback
view) or shown (tutor marking view); and having access to a spreadsheet view of the
marks and sub-marks. All team members commented that this ability of the EPSS to
record marks once on the tutor marking view and then have access to them in the
spreadsheet view was of great value.

The quality assurance of the marking key rubric occurred during a number of unit
team meetings, through individual feedback, and from observation of the tutors by the
researcher during the development cycle. This process was not typical of other DCs in
that it was spread out over about five weeks.

A suggestion voiced by all tutors during this activity (following from Activity 4)
was to reduce the four comment boxes, one under each of the grade descriptor, to one
box covering all the grade descriptors and use the marked radio button to indicate the
grade the student achieved for that criteria. All tutors found that the descriptor box was too small to hold the amount of comments they wanted to give. This design aspect was incorporated into the following version of the student feedback and tutor marking views. However, there was insufficient time due to the drawn-out marking process for the researcher to modify this within the current development cycle (DC1). As the tutors were completing DC1, a number of tutors also began to suggest a reduction in the number of comment boxes, that is, have one box per group of criteria instead of one per criterion. These design features were incorporated into the next two development cycles (DC2 and DC3).

In the student feedback view, the use of the radio buttons without marks showing did work as expected. However, the feedback from the students, via the tutors, was that where four radio buttons were used it was unclear what grade mark they had received. This issue of multiple buttons occurred in the subject knowledge/content criterion within the high distinction grade descriptor, as shown in Figure 6.7. This issue was due to the space within which the grade marks had to be positioned and meant that they had to be placed on two lines. A number of suggestions were discussed during the unit team meetings and individually with the researcher to overcome this problem. Specifically the suggestions were to reduce the number of marks allocated as per other descriptors, to add text to the radio button to indicate the grade mark, to show the mark, or to implement a combination of these options. Time constraints permitted only the reduction in the number of possible grade marks allocated in the following two development cycles (DC2 and DC3). During the teaching break between semesters, the researcher explored a number of possibilities to achieve these design suggestions and some of these were incorporated into the semester two development cycles (DC4, DC5 and DC6). These features are discussed in Chapter 7.

![Figure 6.7 Criterion where four radio buttons were used in a grade descriptor (DC1).](image-url)

The allocation of marks, both criteria and grade descriptor marks, were an ongoing issue of discussion throughout the research study. The need for more marks to
be available at the lower end of the grade range became evident as the tutors began to use the EPSS. The team decided to add two additional marks to the pass grade descriptor: a zero for not attempted and a mark below pass, for not satisfactory or not complete. Time allowed this to be incorporated into the next two development cycles (DC2 and DC3).

The researcher was able to introduce most of the design suggestions into the following two development cycles (DC2 and DC3) despite the time constraints. However, some of these suggestions could not be implemented as DC2 and DC3 overlapped in time: the poster presentation assignment (assignment 2) was due in Week 12 and the essay assignment (assignment 3) was due in Week 13. This placed both the unit team and researcher under a great deal of pressure to complete at times concurrently all the activities of DC2 and DC3.

**Poster presentation development cycle (DC2)**

**Description of assessment task**

The *Poster Presentation* assignment was the second assessment item within the Unit EDL1101 *Learning and Development I*, and was worth 30% of the total unit mark. The sources of data for this development cycle are summarised in Appendix A. This assessment item was designed to allow students to demonstrate their understanding about the way in which people learn, and the theorists who had contributed to this understanding, by using and drawing on the readings and the results of their first assignment (*Oral Presentation*). The students were expected to convey this understanding to the rest of the class through a poster presentation. The detailed guidelines for this assignment are shown in Appendix D.

**Description of activities**

The *Poster Presentation* DC went through a series of activities similar to the previous cycle and is summarised in Table 6.5 below.
Table 6.5

Poster presentation development cycle (DC2)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Contributor(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Marking key design and development</td>
<td>Researcher</td>
<td>Placed criteria headings only and no grade descriptors on the assignment 2 website recording page, and emailed tutors a similar page in a Word document. These criteria headings were based on Assignment 2 guidelines in the unit outline and generic poster presentation rubrics. (Poster Rubrics - (McCullen, n.d.)</td>
</tr>
<tr>
<td>Tutors</td>
<td>Logged to the unit website and entered their comments, judgements, thoughts, etc, under the grade descriptors for each criterion or used the Word document that was emailed to them.</td>
<td></td>
</tr>
<tr>
<td>2: Marking key quality assurance</td>
<td>Coordinator</td>
<td>The online feedback from the tutors and the feedback provided on the Word document was collated and summarised to produce a collated set of criteria and grade descriptors.</td>
</tr>
<tr>
<td>Team</td>
<td>This collated set of criteria and indicators were reviewed and edited during a team meeting. These suggestions and comments were then used to produce the final marking key rubric. The allocation of marks for both criteria and indicators occurred during this activity.</td>
<td></td>
</tr>
<tr>
<td>3 Incorporation of marking key within the EPSS</td>
<td>Researcher</td>
<td>Based on the results of the last meeting where the criteria, grade indicators and sub-marks were decided upon, the completed marking key rubric was incorporated into the EPSS.</td>
</tr>
<tr>
<td>Team</td>
<td>The completed EPSS was presented to the team at the regular face-to-face meeting for a final signing off prior to its use. Both a hard copy of the marking key was handed out for final review and the electronic version was demonstrated.</td>
<td></td>
</tr>
<tr>
<td>4 &amp; 5: Pre-marking moderation and Marking combined</td>
<td>Team</td>
<td>A copy of the blank EPSS containing only the test student development record was placed on all the tutors’ laptops. Before marking started, the coordinator selected a number of sample poster presentations for an on the spot moderation meeting. These samples were usually marked in pairs before the moderation discussion.</td>
</tr>
<tr>
<td>Tutors</td>
<td>All the students poster presentations were displayed around the rooms with only the poster number showing. To mark a new assignment, the tutors created a new tutor marking view and entered the poster number, the marks and comments.</td>
<td></td>
</tr>
<tr>
<td>6: Post-marking moderation</td>
<td>Researcher</td>
<td>When marking was completed, the tutor databases were copied from the tutors’ laptop computers and combined by the researcher to produce an Assessment 2 database.</td>
</tr>
<tr>
<td>Coordinator</td>
<td>This allowed the coordinator to moderate the marks and comments between the tutors and overall for the unit. The spreadsheet view and student feedback sheets were sorted by tutor and student surname, printed and returned to the tutors.</td>
<td></td>
</tr>
</tbody>
</table>

The main differences between DC1 and DC2 were in Activities 1, 2, 4 and 5. In Activities 1 and 2, the tutors were given more responsibility in helping to develop the marking key rubric by being directly involved in the development of the grade descriptors for each criterion, while the coordinator and researcher jointly developed the criteria headings, and the researcher developed the Internet recording page. Activities 4 and 5, pre-marking moderation and marking activities, differed from those of DC1 in that they occurred at the same time, and marking was completed during the evening that the posters were collected, instead of over a period of two weeks.
The initial development of the marking key rubric in Activity 1 was simplified in this DC. The coordinator decided to only supply the initial criteria headings for the rubric and not the grade descriptors for each of the criteria as was done in DC1. This was decided upon in consultation with the researcher and was based on the limited time to develop the grade descriptors, the lack of experience in marking poster presentations by both the coordinator and tutors, and the limited number of sample rubrics that were found to be similar to the assignment objectives during the literature investigation. These initial criteria headings were based on the assessment guidelines, the previous 2002 marking key (Figure 6.8) and generic poster presentation rubrics found at the Poster Rubrics website (Poster Rubrics, 2005). This resulted in the tutors being given the opportunity to contribute more directly in the development of the grade descriptors from the start instead of providing them with suggested grade descriptors as in DC1.

EDL1101 Semester 2, 2002
CRITERIA FOR ASSESSMENT OF POSTERS
1. Catches interest and can be read from a distance of 1 metre
   HD  D  C  P1  P2  F1
2. Conveys the central ideas and provides key details in ways that show an understanding of the concepts involved.
   HD  D  C  P1  P2  F1
3. Appropriate level of detail.
   HD  D  C  P1  P2  F1
Comments:

Figure 6.8 Previous poster marking key (DC2).

Due to the tutors’ Internet access issues encountered in DC1, the coordinator and researcher decided to develop a Word document version of the rubric marking key. The online database and Word document (see poster marking key development document in Appendix E) were made available to the tutors during Activity 1. This allowed the tutors time before the start of Activity 2, quality assurance of the rubric marking key, to record their suggestions and comments, and gave the coordinator time to collate them before the unit team meeting. This collated version was circulated to the tutors during the unit team meeting to record comments and suggestions, and finalise the criteria headings and grade descriptors and sub-marks. Recommendation arising from DC1 resulted in the incorporation of the following innovative features to the marking key:

- The number of grade descriptors in a number of criteria were reduced from four to three, as the unit team felt that the criteria could not be realistically divided further, e.g., the criterion Colour had only three grade descriptors;
• The criteria were grouped under three headings: Presentation, Information and Generic or Objective criteria;
• The comment boxes were reduced from one per descriptor to one per group of criteria headings;
• The grade descriptor marks were again hidden and only the total mark and grade were shown on the student feedback view. Also, two marks were added to the Pass descriptor: 0 and Not Completed. Lastly, the number of possible allocated marks was reduced in the High Distinction grade descriptor from four to two; and
• Negative marks were introduced for objective criteria that were considered either objective in nature or generic, for example grammar and spelling, referencing, poster and text size.

The coordinator and researcher used these recommendations and suggestions from DC1, and the edited and collated comments and suggestions for the rubric to produce the final version of the marking key. The researcher then incorporated this final signed-off copy into the EPSS as part of Activity 3. This process included the development of the student feedback and tutor marking view forms that are shown in Figures 6.9 (student feedback view) and 6.10 (tutor marking view). The tutor marking view had the grade descriptor marks shown and arranged to fit the landscape display of the computer monitor and the tutors’ three comment boxes were placed at the bottom of the three groups of criteria headings. This was done to eliminate or reduce scrolling.
Figure 6.9 Poster presentation student feedback sheet (DC2).
A number of additional features were added to the tutor marking view. Firstly, a pop-up menu showing the marks per criteria (see Figure 6.11) was introduced by the researcher to reduce the amount of cursor movement in anticipation of the marking occurring using laptops without mouse access (that is, standing or sitting without a desk). The radio button method of mark entry was retained, as there was no time for formal training in the new method of mark entry. The researcher considered that adding the pop-up menu increased the tutors’ options, and could be used to adjust the marks more easily after they had been entered. Secondly, a field was added to record the poster presentation number and was displayed on all views.
Figure 6.11 Pop-up menu for recording marks.

Due to the nature of this assessment item, a poster presentation measuring one metre by one metre, it was difficult for the tutors to store, access and mark this assignment over an extended period. The unit team decided to complete the marking of the poster presentations in two marking sessions of about 5 to 6 hours each during two evenings. This meant that pre-marking moderation and marking (Activities 4 and 5) were combined, as the marking of the student poster presentations needed to be completed within 24 hours of them being handed-in and displayed in the classroom.

The researcher, having only developed the EPSS once before, realised that the tutors were all assuming that the tool would work without any problems and that there was no time to correct any glitches in the EPSS if they did occur on the assessment nights. The researcher reasoned if any did occur and could not be quickly fixed, then the tutors would have to record the marks and comments by hand and might, as a result, be less enthusiastic about continuing with the study. Thus, the tool was tested by the researcher as thoroughly as time permitted, so that the tutors could focus on completing the marking process as quickly as possible within the tight time constraints.

Poster presentations were de-identified by allocating a number to each, which the tutor recorded next to the student’s name and number. This allowed students to peer-mark a number of posters by completing a feedback sheet (different to rubric marking key) without knowing the identity of the student. On the first day of marking, four classes were marked involving around 100 posters.

Pre-marking moderation (Activity 4) involved tutors marking a number of posters selected by the coordinator. During this process, a number of tutors stood together marking and commenting on the same poster presentation while recording comments and marks into the EPSS on their laptops. At the completion of this moderation process,
all the tutors came together to discuss their comments and marks and a final consensus was reached.

Due to the tutors’ positive feedback in DC1, all the pre-marking moderation and marking was done with the EPSS on laptops, including one tutor who had never previously used a laptop. Two copies of the EPSS, one blank, the marking moderation version, and the other containing the list of students for that tutor, the tutor marking version, were placed on each tutor’s laptop. Tutors required assistance at times, as the tool was in its developmental and prototype form of production and did not have the refinements expected of a commercial tool. However, as the tool did not need many refinements, it meant that the training needed was limited to the very basic functionality of the tool. Added to this, the tutors were collaborators in the development of the marking key rubric so they were very familiar with the layout, content and purpose of the tool.

The researcher was present during Activities 4 and 5 to observe, record feedback, and to give assistance if needed in the use of the EPSS. The researcher explained and demonstrated how to create a new tutor marking view during the moderation activity, record the poster number, and navigate between the spreadsheet view and the tutor marking view. Specifically, in the tutor marking view, tutors were shown how to enter the poster number (if not already recorded), after having found the student whose poster presentation they were about to mark in the spreadsheet view, and then move to the tutor marking view. Tutors sat on wheeled chairs with laptops on their knees and pushed themselves around marking the poster presentations. The whole process (Activities 4 and 5) took from 5 pm to about 11 pm to complete over two nights.

In the final activity, post-marking moderation, the individual tutor’s EPSS containing their checked and reviewed marks and comments, were copied by the researcher. The copy remaining on the laptops acted as a back-up copy. The researcher then combined the individual tutor’s EPSS to produce a poster presentation EPSS containing all the students’ marks and comments. This master poster EPSS was then copied on to the coordinator’s computer, and enabled the coordinator to moderate for both quality assurance and control. Quality control covered the range of marks and comments that occurred between tutors, and the overall spread of results for the whole unit. The EPSS made this process of moderation easy, as the spreadsheet view of the students’ marks was part of the tool. The tool allowed the coordinator, with the assistance of the researcher, to find or track both individual tutor’s class and unit results,
and then sort these by any field (for example total mark, student name or tutor’s name). As the tool also calculated the grade based on the total mark for the assignment, the coordinator was able to visually see and count the number of grades being allocated by tutorial group and for the whole unit.

The coordinator, having completed the quality control, completed the final step of the moderation process: the quality assurance of the comments and presentation of the student feedback view, checking that all missing student marks were accounted for. The whole unit spreadsheet view was then sorted by student name and a PDF copy saved and printed (this provided a hard copy back-up of the assessment results without the comments). Finally, the student marking sheet view was selected, sorted by tutorial groups, and then by student surname, printed, and placed in tutorial envelopes for tutors to collect and return to the students. This procedure was completed within a week of the poster presentation marking.

**Discussion of findings from DC2**

In this DC, the tutors’ contributed to the development of the marking key. The Word document option for recording comments and suggestions had proved to be more acceptable to the tutors than the online-based option (although the researcher and coordinator first considered using an Excel spreadsheet document, the fact that the tutors were more familiar with Word documents persuaded them to use a Word document). The introduction of an alternative method of recording comments and suggestions was partly motivated by that fact that many of the tutors continued to have problems accessing the Internet from home; most tutors also found recording their comments and suggestions by hand on the Word document more appropriate for this type of moderation exercise.

The researcher found that the tight time constraints of this cycle meant that little time was left for testing and gaining feedback from the coordinator. The modifications of the EPSS to create fields occurred without much difficulty, as these followed the same procedures developed in DC1. However, the display layout in the tutor marking and student feedback views required a great deal of design work to fit on one A4 sheet. The addition of the generic or objective criteria section required also required some additional work.

The moderation process consisting of creating a new record and then just adding the poster number worked well, but the sharing of this marking information between the
tutors in the discussion of results was difficult, and some reverted back to the hard copy of the rubric marking for the moderation process. This would have been an ideal situation to trial the application of wireless technology, as the tutors could all have been working from the same EPSS application. This would have allowed them to have instantly seen what the other tutors had given for the same piece of work. The researcher also observed that very few tutors used the pop-up menu to record the marks and were content with recording them by clicking on the radio buttons. Overall, the tutors found the EPSS performed well and the navigation between the spreadsheet view and tutor marking view worked well. The tutors found the detailed rubric helped greatly in the allocation of marks. Consequently, they only needed guidance and help occasionally from other tutors or from the coordinator, when the mark allocated to the poster was a borderline case.

The collation and combining of the tutors’ individual EPSSs worked efficiently and all was completed within two days of marking. The tutors, with the researcher’s help, spell-checked their comments, and reviewed their marks before they were copied and combined into the master EPSS. The post-marking moderation went extremely smoothly considering the very tight time constraints under which the researcher and coordinator worked. The spread of grades was acceptable according to the University guidelines, and the tutors felt that the range of marks reflected the students’ abilities.

**Essay development cycle (DC3)**

**Description of assessment task**

The Essay Assessment was the third and last assessment of the Unit EDL1101 Learning and Development I. The sources of data for this development cycle are summarised in Appendix A. It consisted of a 1500-word essay and was worth 40% of the total unit mark. To complete this assessment, the students were asked to write an essay in which they were to assert their understanding of the process by which learning occurs, and provide supporting evidence from the reading and research regarding the theories they had elaborated on in their poster presentation. The essay assessment criteria were provided in the unit outline (see Appendix F).
Description of activities

This DC went through seven activities similar to the previous two development cycles (DC1 and DC2), however as this cycle overlapped considerably with DC2, the time available for development was significantly reduced. The activities involved in DC3 are outlined and described in Table 6.6.

Table 6.6
Essay development cycle (DC3)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Contributor(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Marking key design and development</td>
<td>Coordinator and Researcher</td>
<td>Developed the criteria and grade descriptors.</td>
</tr>
<tr>
<td></td>
<td>Tutors</td>
<td>Reviewed the marking key draft and made comments on a Word document supplied as an e-mail attachment.</td>
</tr>
<tr>
<td>2: Marking key quality assurance</td>
<td>Tutors</td>
<td>Brought their annotated copies of the marking key to the moderation meeting. The sub-marks were agreed upon and the grade descriptors were finalised.</td>
</tr>
<tr>
<td></td>
<td>Coordinator</td>
<td>Updated the marking key based on the suggestions and feedback provided at the unit team moderation meeting.</td>
</tr>
<tr>
<td>3: Incorporation of the marking key within the EPSS</td>
<td>Researcher</td>
<td>Incorporated the signed-off marking key into the EPSS.</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>The completed EPSS was presented to the team at a regular meeting for a final signing off prior to its use.</td>
</tr>
<tr>
<td>4: Pre-marking moderation</td>
<td>Coordinator</td>
<td>Selected a number of student sample essays for moderation and gave the tutors copies of two samples and the marking key.</td>
</tr>
<tr>
<td></td>
<td>Tutors</td>
<td>Tutors had three days in which to mark the two supplied student sample essays before the moderation meeting.</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>Moderated their marks and comments on the samples provided and came to a consensus at the moderation meeting.</td>
</tr>
<tr>
<td>5: Marking</td>
<td>Researcher</td>
<td>Divided the EPSS database into tutor groups and copied to tutors’ computers for marking.</td>
</tr>
<tr>
<td></td>
<td>Tutors</td>
<td>Marked the essays using the EPSS.</td>
</tr>
<tr>
<td>6: Post-marking Moderation</td>
<td>Researcher</td>
<td>Combined the individual tutor databases to produce an essay assessment database.</td>
</tr>
<tr>
<td></td>
<td>Coordinator</td>
<td>Carried out moderation, both quality control and assurance of the students marks and feedback. The spreadsheet and student feedback views were sorted, printed and returned to the students.</td>
</tr>
<tr>
<td>7: Management</td>
<td>Researcher</td>
<td>Linked the three EPSSs containing the results of the three assessment results to a master student list. This master student list database displayed the student mark from each assessment item plus the overall mark and grade for the unit.</td>
</tr>
<tr>
<td></td>
<td>Coordinator</td>
<td>This information allowed the coordinator to carry out quality control and assurance of the final total mark for all the students before they were submitted to the examination board.</td>
</tr>
</tbody>
</table>

In Activity 1, the coordinator, in collaboration with the researcher, developed a working copy of the marking key. Limited time was available to carry out Activity 1, as
DC2 was still being completed. The coordinator decided that, due to the shortage of time, the tutors would not be involved in this early activity of the marking key development as had occurred in DC2. The working copy of the marking key was developed using knowledge gained from DC1 and DC2 in terms of structure and content, the previous year’s marking key (see Appendix G), and the University generic grade descriptors. This material was consolidated by the coordinator into a working copy of the marking key and entered into a Word document (see Appendix H). This document was emailed to tutors and placed on the Internet, so that tutors had the opportunity to record comments and make suggestions before the marking key moderation meeting.

Activity 2, the marking key moderation, involved the tutors attending a moderation meeting. The tutors’ recorded comments and suggestions were added to their copies of the working marking keys. Both the overall wording of the marking key and the wording of each grade descriptor within each criterion were discussed vigorously to ensure they reflected the desired improvement in the student’s performance between the grade descriptors. For example, for the criterion Introduction, the grade descriptor text for Credit: *Outlines most aspects of the purpose and topic for discussion*, was amended to *Outlines most aspects of the purpose and/or topic for discussion*. Additionally, the marks per criterion and the break-up of the grade marks were agreed. After the meeting, the coordinator, using the comments and suggestions, finalised the marking key and forwarded a copy to the researcher for incorporation into the EPSS.

During Activity 3 the researcher incorporated the signed-off marking rubric and once again created the three views: *student feedback* view (Figure 6.12), *tutor marking* view (Figure 6.13), and *spreadsheet* view (Figure 6.14). This version of the EPSS incorporated and built upon the features developed in DC2. The criteria were grouped by headings and three comment boxes were added including an overall comment box. Negative marks were again applied to the objective and generic features of the essay (e.g., spelling, grammar and referencing).
The researcher, in discussions with a number of tutors and based on his knowledge of essay marking and their growing understanding of the possibilities of electronic assessment, suggested that a global comment box could be added to the tutor
marking view. This global comment box would allow the tutor to add comments that would appear on every recording page and would be used to hold generic or recurring comments that the tutor could reuse or modify by copying and pasting into the overall comment box (see Figure 6.13). Time did permit this feature to be added to the tutor marking view. Examples of some of the tutors’ global comments were:

- “Remember to discuss the purpose of your essay in your introduction, providing a ‘road map’ for the reader”;
- “You have only drawn from the required reading for the unit. Additional literature has not been used”;
- “A thorough understanding of the topic has been conveyed. Well done”;
- “Check University referencing guide for in text standards”;
- “Your conclusion needs to summarise your essays arguments clearly”; and
- “The theories have been explained and need to be further substantiated from a range of relevant literature”.

To enhance this new feature of the EPSS, a facility was provided to allow tutors to automatically add a pre-set text to the overall comment box by clicking on a button. The trailed text was “I strongly suggest that you make an appointment to see our Faculty Learning Advisor, <name>. Room <number> email <email address>.” (see bottom of Figure 6.13).

Pre-marking moderation, Activity 4, took place over three days, with the coordinator firstly photocopying two sample student essays plus the rubric marking, and giving these out to the tutors, who were requested to mark them before the unit team moderation meeting. At the meeting, a number of minor spelling and grammar errors on the marking key within the EPSS were reported back to the researcher for updating before the tool was divided into the tutor groups for marking. At the end of the moderation meeting, the tutors collected their own student essays for marking, and times were arranged to meet with the researcher to obtain their individual tutor’s EPSS copy.

Activity 5, the marking of the essays by the tutors, occurred either at home or in their offices. As the tutors completed their marking, they arranged to meet the researcher so the researcher could collect their completed database. During these meetings, a spell-check was performed on all their comments and a hard copy of the spreadsheet view sorted by total mark was obtained. The researcher then combined the individual tutor databases to form a master EPSS containing all the students’ marks and comments. This master EPSS was forwarded to the coordinator.
The coordinator, with the assistance of the researcher, carried out the final quality assurance and control (Activity 6). On completion of these tasks, the **student feedback** view was sorted by tutorial group and student surname, and then printed. The printouts were then placed in envelopes by tutorial group and returned to the tutors, who attached them to the student essays and returned them to the students.

As this was the final activity in the final development cycle for the unit, an additional task was required of the EPSS: bringing together the three assignment marks and producing an overall total mark and grade for the unit. The researcher had already anticipated this requirement and had worked on the development of this feature since the completion of DC2. The student number was used to link the three master EPSS databases for each assignment to a master unit EPSS database. The fields containing the total mark for each assignment were brought across and added to produce a grand total mark and grade for the unit (see Figure 6.14). The coordinator was then able to carry out the final quality control of the unit marks and grades before submitting the marks to the Examination Board for final approval.
To help in the process of quality control the EPSS was also used to produce a number of summaries and graphs of the student marks and grades to aid the coordinator judge the acceptability of the range of marks between the tutors and overall for the unit.

**Discussion of findings from DC3**

The Activity 1 process was similar to the one used in DC1, with the coordinator developing a working copy of the marking key with assistance from the researcher. This process worked well overall. The word document emailed to tutors to record comments and suggestions achieved its aim and saved much-needed time.

Although Activity 2 was performed under tight time constraints, the tutors found time to record comments and suggestions on the word document before the marking key moderation team meeting. The coordinator found the collation of comments and suggestions difficult under the tight time constraints. This resulted in a number of minor grammatical and grade descriptor errors being reported in the final version of the marking key during the pre-marking moderation meeting (Activity 4). These minor errors were corrected by the researcher before the tutors received their copy of the EPSS for marking in Activity 5.

The researcher considered that despite the tight time constraints experienced during DC3, the incorporation of the marking key rubric into the EPSS (Activity 3) was performed much more efficiently. This was due to what the researcher had learnt and the procedures developed during DC1 and DC2. As in the previous development cycle (DC2), more features were added based on observations, feedback, and recommendations from the tutors and coordinator.

During Activity 4, the pre-marking moderation, all tutors once more commented on the benefits of the detailed marking key. This detailed marking key allowed the unit team moderation meeting to focus more on the content of the sample student essays and less on the understanding and interpretation of the marking key. Tutors found using the
hard copy of the tutor marking view with the marks showing a satisfactory way to record comments and suggestions before and during the moderation meeting.

The marking process in this DC was a more typical assessment process for the students, tutors and coordinator than had been the case for DC1 and DC2, mainly because the student essays all had the same due date and were marked and returned to the students within two weeks. The feedback from the tutors during the marking process was again positive. They appreciated the global comment box and tended to make use of the auto-add square button field if they needed to.

The negative marking, however, proved too severe for students whose marks were at the lower end of the scale, and somewhat distorted their overall mark. Additionally, as this was a more typical assessment process, an issue arose that had not previously been encountered in DC1 and DC2: the late submission of assignments, which also entailed a penalty mark. A number of ideas were suggested to overcome these problems, and the two solutions that were implemented were respectively marking the few students affected less severely, and creating a field to hold a late penalty mark that could be entered manually. This allowed tutors to adjust their individual tutor marks before they were combined.

The performance of the post-marking moderation (Activity 6) was significantly improved by the features provided by the EPSS. The coordinator particularly praised the ability of the EPSS to bring all the tutors’ results together including the sub-marks and comments.

As this was the final assessment for the semester, DC3 involved an additional management activity (Activity 7), which was also significantly enhanced by the application of the EPSS. The EPSS was used to collate the three assignment marks and produce an overall unit mark and grade for each student. The application of the EPSS to perform this management activity went smoothly and it significantly contributed to the ease and efficiency of the process. For example, the tool allowed a graphical view of the grade allocation by tutor, whole unit, or any other grouping (see Figure 6.15). The use of the EPSS, as the coordinator noted, “removed all of the busy work” that used to be involved in this important process of combining the student assessment marks.
Figure 6.15 Distribution of grades for one tutorial group (DC3).

**Discussion of findings from Semester 1**

This section discusses the overall findings from the three collaborative DCs in Semester 1, and includes the findings that occurred within the DC activities and the iterative changes to the layout and features of the EPSS. The findings were based on the observations made by the researcher during each DC, and on feedback provided during the unit team meetings and obtained from individual tutors while they were using the EPSS. In addition, at the final unit team meeting, the researcher organised a brief group feedback session to capture any comments, suggestions, and other feedback that might not have been captured during the semester, and to thank them for their participation and bring closure to this part of the study.

**Findings from the development cycle activities**

The development of assessment-specific marking key rubrics proved more difficult and demanding than the researcher or the coordinator had anticipated. The marking key samples obtained from the literature search proved useful to an extent, in that they provided an overall structure and an indication of the appropriate language to use in the development of the grade descriptors. Significantly, the researcher observed that the cognitive load and time required by the team members to develop the detailed marking key decreased over time, as team members were able to build upon the skills and knowledge obtained while developing the previous marking keys.

The use of the Internet in Activities 2 and 4 proved to be partially successful as only campus-based tutors were able to access and record their comments and suggestions. To solve this Internet access issue, the researcher and coordinator, in
consultation with the tutors, explored other methods of capturing the tutors’ comments relating to Activities 2 and 4, and eventually the researcher developed a word document that was emailed to the tutors, who could either print it or complete it electronically. This alternative process was effective and was well received by the tutors.

The marking key quality assurance (Activity 2) proved most beneficial and useful for the tutors, as they were able to gain a degree of understanding and took ownership of the marking key. The tutors were less familiar with this type of moderation process and commented that they had not previously been involved so early in the assessment process. They felt that this early involvement in developing the marking key helped in the quality control of the students’ marks later in the assessment process. Being involved in the quality assurance of the marking key gave tutors a much deeper understanding of what the marking process was trying to achieve and what specific words in the marking key meant in relationship to the specific assessment task. As one tutor commented: “the process used was ideal in establishing the criteria clearly. It is essential that the whole team has input at this activity to make sure all understandings are clear.”

Another tutor commented that this moderation process “was ideal in that I could first think about it at home then bring my suggestions and ideas to the group”. While another said, “it is useful to discuss understandings with other markers and construct the rubric together, this clarifies my understanding of what is meant by each category.” However, another tutor felt that the rubric “needs to be revisited and changed during marking, if possible, because that’s when the weaknesses become apparent.”

The coordinator found that the time and effort required in editing and collating the tutors’ comments and suggestions, although significant, particularly in DC1, proved both useful and beneficial when viewed over the whole assessment process. The coordinator and tutors all commented that this time and effort spent in Activity 2 was largely compensated by the efficient, smooth-flowing and more consistent allocation of marks in Activities 4, 5 and 6. The coordinator and tutors commented that these activities at the end of the task assessment process were completed in less time and with less stress than in previous assessment processes used in other units.

Activity 3, the incorporation of the marking key rubric into the EPSS, was a just-in-time development process for the researcher and had to be completed usually within a week. The researcher gained efficiency as the cycles progressed, but as new or modified features were continually being added, this efficiency was often lost in
determining how to add features to the EPSS or modify existing ones. Because no time was available for training as new features were added to the tool, these were explained and demonstrated at the unit team meeting and individually on a needs basis by the researcher. This approach, although not ideal, proved very effective, partly due to the tutors’ confidence that the EPSS would perform as demonstrated and their willingness to seek help and guidance in using the EPSS from the researcher and other tutors.

Overall, the pre-marking moderation (Activity 4) was a success, but because not all the tutors had laptops, this process could not be accomplished just using the EPSS. The tutor marking view was used to mark student samples of work selected by the coordinator. These samples were marked either before the unit team meeting or during the meeting, and the marks and comments were discussed and moderated. The use of the videos of the oral presentations in the pre-marking moderation activity in DC2 also proved to be a successful strategy.

The actual marking with the EPSS took place in Activity 5. Although most tutors were either novices or advanced beginners in the use of database applications, they found the recording of students’ marks and comments an easy and efficient process using the EPSS. This reported ease of use was partly because the EPSS student feedback view was an exact copy of the marking key developed on paper, and the tutor marking view was very similar, with only minor layout changes, while the spreadsheet view looked similar to a typical spreadsheet document. The problems the tutors did encounter were the typical issues associated with the features of a database: multiple views, auto-save, find and sort, and printout functions.

One tutor commented on Activity 5 of DC1, noting that the EPSS

\[ \ldots \text{streamlined the marking, made it more consistent and therefore fairer. It helped me focus on the important points.} \ldots \text{I found I couldn’t mark the presentations on screen during the actual presentation. This is because I was getting used to my new laptop, getting used to the on screen marking and the students. I think I would use it during the presentations next time.} \]

While another commented during the final development cycle that the EPSS was:

\[ \text{Very manageable, you could mark wherever you were comfortable using the laptop. I liked the [global] comment box at the bottom that you could draw from.} \]

Coordinator and tutors consistently commented throughout the three DCs that the time and effort spent developing the marking key was worthwhile, and largely offset by the final Activities, 4, 5 and 6 being completed more efficiently in terms of time, effort
and satisfaction. Additionally, the tutors commented that they felt more confident in their marking. Overall, the coordinator and tutors felt that the detailed marking key containing the grade descriptors was an essential feature of the tool’s success.

Activity 6, the post-marking moderation, was a success as little or no mark correction was required. The distribution of marks and grades between tutors, tutorial groups and within the whole unit for each assignment were well within the acceptable university guidelines. The process of bringing the individual tutor EPSS containing their student marks and comments together into a master assignment EPSS proved to be a quick and easy process.

The management activity (Activity 7), involved collating the three assessment marks and providing a final unit mark and grade. This proved to be a process easily accomplished by the EPSS. The coordinator found using the EPSS much more straightforward and much less complex than previous methods used to collect and collate marks from tutors.

**The evolution of the layouts and content of the EPSS**

The main views of the EPSS were the student feedback, tutor marking and spreadsheet views. Although the number of views did not change over the three DCs, their layout and content changed greatly. The changes to the layout and content of each of the three views are discussed below.

The student feedback view was designed to be printed as an A4 portrait document. During the final meeting of the semester, some tutors commented that a number of students had found that they did not understand how the sub-marks were allocated and what they stood for, and that the overall mark did not provide enough feedback for them. A similar issue had been raised by some tutors during DC1 that related to the fact that there were not enough sub-marks at the Pass grade level and there were too many for the High Distinction level. This issue had been progressively resolved over the DCs. The unit team felt that the issue raised by the students was a valid issue, but still decided not to show the sub-marks on the student feedback view, as students had argued over sub-marks in the past and the unit team wanted the students to focus on the grade descriptors rather than on the actual sub-mark.

The researcher, having gained a greater understanding of the possibilities of the software used to develop the EPSS, suggested that instead of hiding the sub-marks, the actual grade words could be used instead (see Figure 6.16). The EPSS would
automatically translate the word mark into a numeric mark. In addition, the sub-marks allocated per criterion would be based on the mid-point of the grade being allocated, that is, once the criteria marks were allocated the sub-marks would be calculated using a formula to determine the mid-point of the grade. For example, if the criterion was marked out of 10 marks, then a Pass would be 5.5, or the mid-point between 50% and 60% (range for a pass out of 100%), a Credit would be 6.5, a Distinction would be 7.5 and a High Distinction would be 9. Also, at the Pass grade level two sub-marks were added: 0 for Not Attempted and 4 for Not Complete. In addition, where the criteria did not warrant a four-grade division, a three-grade division was used. These recommendations were implemented in DC4 and are discussed in the next chapter (Chapter 7).

Figure 6.16 Grade descriptor marks recorded as text.

Another issue raised by tutors throughout the semester was the excessive number of comment boxes, and the consensus reached was that the maximum number would be one per group of criteria, with the minimum being one.

The tutor marking view allowed tutors to record sub-marks, usually by clicking on radio buttons on the computer screen, and to record comments by typing them into comment boxes on the screen. The use of a pop-up menu to record marks was trialed in DC2 but was abandoned, as the tutors were unfamiliar with this type of data entry. The use of the global comment box and auto-entry of text was a collaborative design feature introduced during DC3. This feature was favourably commented upon and used by all the tutors.

The features of the spreadsheet view did not change over the three DCs. However, the researcher did explore the development of summary views and the exporting of data into Excel to quickly find out how many students had obtained the different grades at the unit and tutorial levels. The main disadvantage of these possible options was the need to be more than novice users of either the EPSS programme or Excel. This meant that training would have been required, which was not within the scope of the study. However, if the feature was deemed useful enough by the unit team, the researcher felt that the possibility did exist to reduce the degree of difficult so that
these features could be incorporated into the EPSS if time permitted. The tutors were comfortable sorting by total mark and counting off the screen the number of students receiving the different grades. They found that due to the processes involved in the marking key development, there was very little need to change students’ marks, and therefore, the sorting and viewing of the grades was just a final check. This was similar in the case of the coordinator, although it was somewhat more difficult to do a quick screen count due to the number of students involved. The coordinator, however, did not find that this was a significant issue.

A number of additional views were developed by the researcher as a result of feedback received throughout the DCs in Semester 1. These additional views were connected to the tutor marking view and enhanced the quality assurance process. These additional tutor views were:

- The spell-checking view, developed to facilitate the spell-checking of tutors’ comments; and
- The comment and spreadsheet view, developed to facilitate the editing of comments and the review of multiple comments on one screen or printout.

The spell-checking view worked like the spell-checker in Word. This view was developed because when using the spell-checker in the tutor marking view, it would detect as a spelling error all the students’ and tutors’ names. The researcher quickly overcame this issue by modifying the tutor marking and creating a new view containing only the student number and the comment box.

The comment and spreadsheet view built on the spreadsheet view by adding the comment text box or boxes below the line of sub-marks. Although this view was not often used, it did prove useful when used by the tutors or the coordinator. This view allowed the tutor to edit, refine and compare comments between a number of students while seeing their sub-marks, total marks and grades on the same screen.

The next chapter (Chapter 7) continues with the discussion of the Development Phase, and covers Semester 2 involving unit EDL1201 Learning and Development II.
CHAPTER SEVEN

DEVELOPMENT PHASE - SEMESTER II

DESCRIPTION AND FINDINGS

Introduction

This chapter continues from the previous chapter (Chapter 6) and describes and discusses the findings from Semester 2 of the Development Phase. The description of each development cycle (DC) follows a similar structure to the one used to discuss the findings of Semester 1, including the description of the task assessment, description of activities and discussion of findings. The chapter concludes with a brief summary of the formative results of the Development Phase.

Due to the performance improvement achieved through the application of the EPSS in the three DCs in Semester 1, the coordinator was keen to continue her involvement in the study. The unit selected was unit EDL1201 Learning and Development II, part of the Kindergarten through Primary Course, with 215 students enrolled. In addition, two of four tutors involved in Semester 1 were part of the unit team for EDL1201, and were also keen to continue their involvement in the study. The two tutors who were new to the study consented to participate in the study and already knew the researcher from his work as IT Support Officer. An additional bonus to selecting this unit for the study was that the same students were involved and they were familiar with the innovations to the task assessment process.

Based on the feedback and findings from Semester 1, the same weekly meetings were held in Semester 2, and similar development activities were followed. However, as will be shown in this chapter, more innovative features and approaches were explored and investigated. These included student involvement in the development of the marking key and in peer group marking. Some of the innovations built upon the feedback received from the tutors and coordinator in Semester 1, while others grew out of the unit team’s expanding experience and understanding of the possibilities that the EPSS offered, not only in terms of enhancing their own performance, but also in terms of facilitating student involvement in the task assessment process. These expanding innovations proved to be a two-edged sword for both the study and the researcher. The researcher needed to find a balance between refining the EPSS and exploring new
innovative possibilities, although this balance often moved towards the innovative possibilities due to the enthusiasm of the unit team members. In each of the three assessment items, different innovations were explored and implemented.

During the unit team’s first meeting for the semester, the researcher handed out the ethics consent forms to the two new tutors and provided a brief introduction to the research study. The researcher also provided a brief summary of the overall findings and results from Semester 1 and answered any questions. Finally, the researcher again emphasised that the unit team’s collaboration and participation was essential to the success of the research study.

**Tutorial paper development cycle (DC4)**

**Description of assessment task**

A *Tutorial Paper* worth 20% of the total unit mark was the first assessment item for the semester. The sources of data for this development cycle are summarised in Appendix A. The assessment task consisted of writing a 1,000-word report. The report would be peer group assessed in the fifth week tutorial session, and the peer group assessment marks would then be moderated by their tutor. The assessment criteria for the tutorial paper are shown in Figure 7.1 and were the same as the previous year (the task requirements are shown in the Appendix I).

<table>
<thead>
<tr>
<th>Tutorial Papers will be allocated marks according to the following criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child A perspective</strong></td>
</tr>
<tr>
<td>The learning event is described clearly</td>
</tr>
<tr>
<td>The social/emotional aspects of the situation are explicit</td>
</tr>
<tr>
<td><strong>Child B perspective</strong></td>
</tr>
<tr>
<td>A different perspective is articulated</td>
</tr>
<tr>
<td>The perspective is plausible in conjunction with that described for child A</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
</tr>
<tr>
<td>Understanding regarding social/emotional aspects of learning is demonstrated</td>
</tr>
<tr>
<td>Effective links with theoretical perspectives are made</td>
</tr>
<tr>
<td>Evidence of reading beyond the set text is demonstrated</td>
</tr>
<tr>
<td><strong>Writing conventions and strategies</strong></td>
</tr>
<tr>
<td>Word count</td>
</tr>
<tr>
<td>Referencing conventions appropriate</td>
</tr>
<tr>
<td>Grammar, spelling, punctuation</td>
</tr>
</tbody>
</table>
Description of activities

At the start of DC4, the coordinator decided, in collaboration with the researcher and the tutors, to explore the feasibility of bringing students into both the development of the marking key (Activity 1) and the marking process (Activity 5). The two tutors who had participated in the previous DCs were enthusiastic, while the two new tutors thought such a learning experience in the task assessment process was valid and useful. However, the researcher was unaware of what difficulties and problems might be encountered in developing these versions of the EPSS. The researcher did realise this was an important area in which the EPSS could be applied, and thus agreed to this innovative use of the technology. This student involvement in the task assessment process meant that two additional activities were incorporated into the DC, development of grade descriptors (Activity 1B) and peer group marking (Activity 5A). A brief description of all activities included in DC4 is given in Table 7.1.

The marking key design (Activity 1A) involved the development of the criteria headings only. These criteria headings were developed through a moderation process during a regular unit team meeting using the assignment requirements and the previous years marking key. The six criteria headings developed were:

- The learning event is described clearly;
- The social/emotional aspects of the situation are explicit;
- A different perspective is articulated;
- Understanding regarding social/emotional aspects of learning is demonstrated;
- Effective links with theoretical perspectives are made; and
- Evidence of reading beyond the set text is demonstrated.

In Activity 1B, development of grade descriptors, the coordinator, in collaboration with the researcher, decided to return to using the online recording process developed in Semester 1 to record tutors’ comments for the grade descriptors. Although the online recording of tutors’ comments and suggestions during the Semester 1 DCs had not worked satisfactorily, it had demonstrated that the process could work. As the problem experienced was mainly due to off campus access issues, it was decided to access the online components of the EPSS within the University Intranet.
The researcher built upon the experience and knowledge gained in the three previous DCs to implement this online recording process. A major logistical issue was encountered: the server used to host the website and EPSS for online access was located on another campus approximately 30 kilometres away and some of the implementation, adjustments and settings could not be done remotely. This meant that, due to the critical time constraints of data entry, on a number of occasions the researcher worked on the server at 6 am to test and implement final adjustments ready for the students to enter data during the 9 am tutorial at the other campus. Although this process was complex and not the ideal method of implementation, the recording process worked smoothly and satisfactorily.

Table 7.1

<table>
<thead>
<tr>
<th>Activities</th>
<th>Contributor(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A: Marking key design</td>
<td>Coordinator</td>
<td>Printed assignment requirements and the previous year’s marking key to be handed out at the unit team meeting.</td>
</tr>
<tr>
<td></td>
<td>Tutors</td>
<td>Based on these handouts, developed the criteria headings during the regular weekly team meetings.</td>
</tr>
<tr>
<td>1B: Development of grade descriptors</td>
<td>Researcher</td>
<td>Added the developed criteria to the EPSS with the grade descriptor boxes empty for recording of student group grade descriptors.</td>
</tr>
<tr>
<td></td>
<td>Students</td>
<td>In each tutorial, groups of five students developed grade descriptors for one the criteria, based in part on the assignment guidelines and the University generic grade descriptors. Each student group then recorded their grade descriptors via the Intranet into the EPSS.</td>
</tr>
<tr>
<td>2: Marking key quality assurance</td>
<td>Coordinator</td>
<td>Collated and edited electronically the grade descriptors developed by the student groups.</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>Further edited the collated grade descriptor during the regular unit team meeting and decided upon a final version of the marking key; published a pdf version of the marking key on the Internet; agreed the criteria marks and grade descriptor sub-marks.</td>
</tr>
<tr>
<td>3: Incorporation of marking key within the EPSS</td>
<td>Researcher</td>
<td>Completed the EPSS based on the signed off marking key. Two versions of the EPSS were developed one for peer group marking and the other for tutor marking.</td>
</tr>
<tr>
<td>4: Pre-marking moderation</td>
<td>Coordinator</td>
<td>Copied a sample range of papers for moderation and gave them to tutors for marking before the unit team meeting.</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>Marked the sample papers individually prior to the team meeting and then moderated the papers, through discussion, during the meeting using the marking key.</td>
</tr>
<tr>
<td>5A: Peer group Marking</td>
<td>Team</td>
<td>Decided how best to manage the student peer assessment process.</td>
</tr>
<tr>
<td></td>
<td>Students</td>
<td>Formed groups of three and marked three assigned papers individually, recording the marks and comments; moderated each essay as a group, recording the results onto a master marking key; recorded these moderated grades and comments on the Intranet.</td>
</tr>
<tr>
<td>5B: Marking</td>
<td>Researcher</td>
<td>Gave each tutor a copy of the student peer group EPSS and the tutor’s marking copy of the EPSS. These two EPSS were linked by essay number. This permitted the tutor to see the student peer group mark and comment as they recorded their grades and comments.</td>
</tr>
<tr>
<td></td>
<td>Tutor</td>
<td>Marked using the EPSS, which allowed them to either mark blind and then moderate, or seeing the student peer group sub-grades and comments.</td>
</tr>
</tbody>
</table>
The tutorial activity of recording grade descriptors involved six groups of five students per tutorial. Each group took one criterion and, through a moderation process, developed possible grade descriptors for that criterion. The assessment criteria and generic grade descriptors that were published in the unit outline booklet were used to help guide the student groups in discussing and developing the grade descriptors. These student-developed grade descriptors were then recorded online by one student from each group at the end of the tutorial. The tutors also collected the paper copies as a backup of the groups’ moderated grade descriptors.

Activity 2, marking key moderation (quality assurance), firstly involved the coordinator in editing the student groups’ grade descriptors that were recorded into the EPSS. An example of two of the groups’ grade descriptors for the criterion social/emotional aspects of the situation are explicit with the coordinator’s edited version is shown in Table 7.2. At the completion of this editing process, the coordinator emailed this working copy of the grade descriptors to the tutors for their comments and suggestions. At a moderation unit team meeting, these edited grade descriptors were discussed and finalised. In addition, during this meeting, the criteria marks, generic marking wording and negative mark allocation were decided upon and finalised. The coordinator then produced the final marking key.
Table 7.2
Tutorial groups grade descriptors with coordinator’s revised version (DC4)

<table>
<thead>
<tr>
<th>Criterion: The social/emotional aspects of the situation are explicit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tutorial Group</strong></td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>Coordinator</td>
</tr>
</tbody>
</table>

In Activity 3, the researcher incorporated the signed off copy of the marking key into the EPSS to produce the now standard three views: the *student feedback*, *tutor marking* and *spreadsheet* views. A number of additional views were created to allow student groups to record their suggested grade descriptors (Activity 1B) and the peer group marking (Activity 5A) via the Intranet, and to allow tutors to view these student peer group marks and comments when recording their moderated marks and comments. The following is a list of all the views created in Activity 3:

- *Student feedback* view (Figure 7.2)
- *Peer group marking* view (Figure 7.3)
- *Blind tutor marking* view
- *Tutor moderation marking* view (Figure 7.4)
- *Spreadsheet* view
- *Spreadsheet comments* view (Figure 7.5)
- *Spell-checking* view

Once the above views had been finalised, two identical versions of the EPSS were produced. One version was for the Intranet for the online recording of the peer group marks and used only the *peer group marking* view (see Figure 7.3). The other version was used by the tutors for marking and contained student and tutorial details.
The researcher linked the student peer group version of the EPSS to the tutor’s by using the combined fields of Essay Number, Tutor Name and Tutorial Time into one unique field called ID. This was necessary because the essay numbers used by the tutors...
were not unique and the method of linking the two databases had been overlooked due to the limited contact time with tutors and focus on the recording process. This linking allowed the tutor to see the student peer group grade marks, total, overall grade and comments, as they moderated and recorded their marks and comments (see Figure 7.4).

The peer group marking view (Figure 7.3) was used by the groups to record their marks and comments via the Intranet. An identical view, blind tutor marking view, was also available to the tutor in their copy of the EPSS so they could, if they wanted, mark blind, that is, without first seeing the peer group sub-marks and comments. The tutor moderation marking view (Figure 7.4) combined the peer group marking view with that of that of tutor’s blind tutor marking view. This combined moderation view was designed for on-screen marking and contained a global field to hold generic tutor comments.

![Figure 7.3 Peer group marking view (DC4).](image-url)
The student feedback view was almost identical to the tutor moderation marking view, except that it was designed for portrait printing and was used as the feedback sheet. This view was different to the feedback views developed in the previous DCs, in that only the criteria headings, grades and total mark were shown, but not the grade descriptors. This was done for two reasons: firstly, the students already had a number of copies of the marking key, and secondly, the addition of the grade descriptors would have made it difficult to maintain the one A4 page limit for printing.

The spell-checking view was once again used to spell-check the tutors’ comments, and contained only the comment and student number fields. The spreadsheet comments view was a modified view of the spreadsheet view and contained the comments of the tutor and peer group and the peer group sub-marks (see Figure 7.5). This summary view allowed the tutor to review the spread of criteria marks and comments between themselves and the peer group.
In Activity 4, pre-marking moderation, a sample range of essays were selected by the coordinator, and photocopied and given to the tutors for marking prior to the moderation meeting. The tutors used blank marking keys to record their marks and comments. These were then moderated at a unit team meeting using their marking key comments and marks to aid in the discussion.

As in Activity 1, Activity 5 was divided into two parts: Activity 5A involved the student peer group recording their marks and comments via the Intranet at the end of the two-hour tutorial session. The peer group marking process had been developed during the unit team meeting and involved the tutor removing the essay cover sheet and any other identifying information and recording a number on the essay before the tutorial started. This essay number had previously been generated by the tutor and written next to the student’s name on the tutorial class list.

At the start of the tutorial, the students were given the first 15 minutes to read as many essay papers as possible to give them an idea of the range and quality of the essays. The tutor then formed the students into groups of three, and each member of the
group marked three assigned papers individually, recording the marks on their blank detailed marking keys. Then, as a group, they moderated their marks and comments for each essay, recording the results onto a master marking key, which was attached to the essay. A student from each group then recorded these results into the EPSS via the Intranet using the peer group marking view (see Figure 7.3). At the end of the tutorial, the tutors collected all the essays including the master marking keys.

Activity 5B, tutor marking, involved the researcher giving each tutor a copy of the student peer group EPSS linked to the tutor EPSS. The researcher showed and explained to the tutors how to open and use the additional features of the tutor EPSS. This included the shortcut from the spreadsheet view to the tutor moderation marking view via clicking on a button next to the student ID number.

Activity 6, the post-marking moderation and re-submission, involved the combining by the researcher of the tutors’ EPSS into a master EPSS containing all the sub-marks and comments of all the students for the assessment task. The coordinator then carried out quality control and assurance of the material using the built-in features of the EPSS. The quality control involved looking at the overall spread of marks within tutorial groups, between tutorial groups and within the unit. Students that had failed were given the opportunity to resubmit with the possibility of gaining a pass grade. The quality assurance process involved checking the comments for educative content, grammatical errors and spelling. The final step involved sorting the student feedback views into tutorial groups and then by student surname, and then printing them for return to the students.

During this DC, the coordinator had been interested in obtaining student feedback about their attitudes towards developing the grade descriptors and the student peer group marking process. An electronic survey was considered an appropriate means of gathering this information. As the unit was already using Blackboard, a learning management system that contained a built-in electronic surveys and testing system, the coordinator, in collaboration with the researcher, decided to conduct a brief survey. The survey was designed to capture the students’ attitudes to being directly involved in the task assessment process, more specifically the development of grade descriptors (Activity 1B) and peer-group marking (Activity 5A). The following section discusses the findings from the student survey.
Student response to Activities 1B and 5A

A brief online survey was developed comprising four questions using a Likert scale, with space for comments after each question. This survey was administered just after the students had completed the final tutorial peer group marking activity. The questions included in the survey were the following: 1) How did you find developing the criteria indicators? 2) How did you find reading and assessing the papers? 3) How did you find the group moderation process? and 4) Do you consider peer assessment a valid and reliable form of assessment? The Likert scale had a range of five possible responses, from 1 to 5 (1 being ‘not useful’ and 5 being ‘very useful’).

A total of 61 students responded to this anonymous online survey, or 29% of the 208 students enrolled in the unit. For the purposes of the survey, all questions scoring 4 or 5 were considered a positive response. The percentage results of the combined positive responses are shown in Table 7.3. The results were encouraging and showed that the students appreciated being directly involved in the task assessment process.

Table 7.3
Positive responses to the e-survey (DC4)

<table>
<thead>
<tr>
<th>Questions</th>
<th>% of combined positive responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 – How did you find developing the criteria indicators?</td>
<td>67</td>
</tr>
<tr>
<td>Q2 – How did you find reading and assessing the papers?</td>
<td>79</td>
</tr>
<tr>
<td>Q3 – How did you find the group moderation process?</td>
<td>75</td>
</tr>
<tr>
<td>Q4 – Do you consider peer assessment a valid and reliable form of assessment?</td>
<td>37</td>
</tr>
</tbody>
</table>

The following is a brief summary of the findings of the survey results, based on the grading of each question and the comments made by students under each question. The first question covered the development of grade descriptors, while the final three questions covered the peer-group assessment activity.

To the first question of How did you find developing the criteria indicators?, 67% of the students agreed that the activity was useful or very useful. This activity endeavoured to encourage the students to think about the task, how it was going to be assessed and the processes involved in assessment. This activity was also designed to empower students by giving them a sense of involvement in the assessment process. Both the quantitative findings and the students’ comments below indicate that this occurred:

*Made you think about the criteria you were to write about and the overall assignment. [Student A]*
It was useful and gave use a better insight and understanding of assessment. Develop a rubric assess by that standard and one hopefully cannot go wrong. [Student B]

How difficult it was to set the marking key. [Student C]

We found this task useful. ... It was a very difficult and provoked a lot of discussion in the groups. [Student D]

Gives a good insight into what is required as well as a feeling of involvement in the decision making process. [Student E]

To the second question of How did you find reading and assessing the papers?, 79% of the students responded either useful or very useful. The students comments on this activity indicated that they found the activity of assessing and seeing how others had answered the same task educative. They indicated in their comments that they learnt from reading other students essays, and many mentioned the fact that the experience showed them how difficult the assessment process is. The following is a sample of comments provided on this question:

Great to be able to compare and see how many different ways the assignment was done. [Student F]

Great to have the opportunity to read other papers. [Student G]

It was good to read others perspectives and to practice this form of assessment. [Student H]

I found this interesting as I was able to compare what I had written, and I realised I had left some things out, and I understood different theorists more clearly from reading peoples conclusions. [Student I]

Assessing! Gave me an idea of the difficulties facing teachers and lecturers with assessing papers with no ‘right’ and ‘wrong’ answers. [Student J]

To the third question of How did you find the group moderation process?, 75% of the students responded either useful or very useful. The actual moderation experience was found by the tutors to improve student understanding of both what is involved in the moderation process and how different and important other people’s views and opinions are. The following students’ comments support this view:

We were able to see the essay from other peoples point of view. We could understand why some people marked harder than others. [Student K]

We learnt to agree to disagree ... it was great to share, communicate and evaluate and even debate (nicely) our reasons for the given grades. [Student L]
The moderation, allowed me to see things in the text that I had missed, I found it useful in seeing other peoples perspectives. [Student M]

To the final question, Do you consider peer assessment a valid and reliable form of assessment?, only 37% of the students responded with either useful or very useful. In reviewing the comments, this low positive response appeared to be due to the students’ lack of experience both in the content area and in the marking and moderation processes. However, all the comments did show that the students had reflected thoughtfully on this aspect of the activity, as the following quotes indicate:

Valid learning experience however at this stage of our course not a reliable form of assessment, prefers lecturers final assessment. Introduction of this [activity] has been a positive contribution to our assessment skills. [Student N]

The process was great because it gave insight into the tutors work and how difficult [it is]. [Student O]

I think it is good to provide feedback, but the ability to assess someone requires experience and further knowledge than what we have. [Student P]

Discussion of findings from DC4

The development of the criteria for this marking key (Activity 1A), worked well and was completed during one unit team meeting. The unit handbook and previous year’s marking key provided enough material for the team to complete the task of developing the criteria headings for the marking key.

The development of the grade descriptors by student peer groups (Activity 1B) had not been tried before and proved very successful. The researcher found that the process of placing these criteria headings on the Intranet for the student peer groups to record their suggested grade descriptors worked well. The students had no difficulties in accessing the Intranet within the tutorial rooms, as each of these rooms had five computers with network and Internet access. Without these facilities, this activity would have been difficult to carry out. The students’ feedback to the tutors was positive, and the involvement of the students in the actual development of the grade descriptors for the marking key was an innovative feature that was facilitated by the use of the EPSS.

The coordinator found that Activity 2 was time-consuming, and that it was difficult to collate and summarise the nine student group grade descriptors per criterion. However, the coordinator found it was a rewarding experience and increased the authenticity of the final grade descriptors. The coordinator also commented that many of the groups found it difficult to develop a progression in performance between the
grade descriptions. Commenting on this difficulty, the tutors felt that the students would benefit from learning about how a moderation process should work. Overall, all tutors felt that Activities 1B and 2 were a success.

The incorporation of the marking key into the EPSS (Activity 3) proved to be a much more efficient process for the researcher, as this was the fourth time this process had been performed. Many of the generic design features of the marking key had now been implemented and what remained was just fine-tuning the design and replacing the content. However, time was needed to design the new views and to link the peer group marks and the tutors’ marks EPSSs. The latter required the creation of a number of additional fields. The researcher considered that next time a unique number, for example the student ID number, could be used for the essay, making the linking of the two EPSSs more efficient.

The student peer group marking (Activity 5A) proved a success. The recording of the sub-marks and comments proved to be efficient and easy for students to do via the Intranet. A limitation was that the students could only record once and not edit their recordings after submitting; if they made a mistake they had to create a new record and make a note in the comment box that this was the one to be used for marking. In reality, very few students needed to repeat the process.

A problem that did arise, although fortunately it only affected a few groups, was that some students recorded the wrong essay number and sometimes tutors misrecorded the essay number against the correct student name. These problems were greatly reduced by the fact that the link between the EPSSs was based on the combined three fields of Essay Number, Tutor and Tutorial Time in the field ID. A number of solutions and suggestions were discussed during unit team meetings and individual tutor discussions with the researcher. One suggestion was “next time to have the EPSS generate a unique number for the essay”, or to “use the student ID number”.

The marking (Activity 5) went well with all the tutors commenting that the ability to see the peer group marks and comments was both a useful and rewarding educational experience. All tutors found that, in most instances, the student peer group sub-marks were either the same as theirs or more severe; and where comments were recorded by students, these were usually more critical than the tutors’. The process of separating the EPSS into tutor groups and distributing these to the tutors worked extremely smoothly, in part because it was the fourth time for some of the tutors, and this meant that they
were able to help the others with any difficulties, thus lessening the load for the researcher.

The post-marking moderation and re-submission (Activity 6) followed similar processes developed in the previous three cycles. These processes again proved to be efficient and productive for the coordinator, who was able to carry out both quality control and assurance in a very timely and rewarding manner. The tutor moderation marking view allowed the tutor to see both the tutor and peer group sub-marks and comments but only the tutor’s total mark was shown. The coordinator, in collaboration with the researcher, developed a number of additional views during Activity 6 that allowed the anecdotal feedback from the tutors that the student peer groups tended to mark harder and commented more severely than the tutors to be investigated.

**Event analysis map development cycle (DC5)**

**Description of assessment task**

The Event Analysis Map was the second assignment in the unit and was worth 30% of the total unit mark. The sources of data for this development cycle are summarised in Appendix A. This assignment required the students to identify, present and analyse a typical learning event that they had observed during their practicum. The assessment (see instructions in Appendix J) consisted of two parts: 1) an event analysis map on an A3 sheet used to present the information, which had to show the relationship between the situation presented and the learning occurring in the situation; and 2) an oral presentation to a small group of students during a tutorial session. The previous year’s assessment criteria shown in Figure 7.6 were used initially to develop the marking key. However, this provided only criteria headings with overall indicators or descriptors, but no breakdown to help tutors with the allocation of sub-marks.
Description of activities

The coordinator, tutors and students had found the process of students recording results on the Intranet efficient and rewarding in DC4. As a consequence, the coordinator, in collaboration with the tutors, decided to continue using this process to record the peer group sub-marks and comments for the oral presentation part of the event analysis map assignment, while the tutors marked and commented on the event analysis map part. However, in this case, the unit team decided to use the peer group mark of the oral presentation as part of the assignment mark, and not just for moderation and reflection by the tutors as had been the case in the previous cycle (DC4). A summary of the activities involved in DC5 is given in Table 7.4.

The development of the marking key (Activity 1) and the moderation of the marking key (Activity 2) were combined in this DC. As this assignment combined aspects of a poster and an oral presentation, the coordinator, in collaboration with the tutors, decided to use the two marking keys developed in the Semester 1 as the starting point for the development of this marking key. This meant that less time and effort would be required. These two marking keys were emailed to the tutors prior to the moderation meeting.
### Table 7.4

**Event analysis map development cycle (DC5)**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Contributor(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2: Design, development and quality assurance of marking key</td>
<td>Coordinator</td>
<td>The oral and poster presentation marking keys developed in Semester 1 were used as the bases for the design and development of the marking at moderation unit team.</td>
</tr>
<tr>
<td></td>
<td>Tutors</td>
<td>This marking key was e-mailed to the tutors for their comments and suggestions. These comments and suggestions formed the bases for the quality assurance of the marking key meeting.</td>
</tr>
<tr>
<td>3: Incorporation of marking key within the EPSS</td>
<td>Researcher</td>
<td>Incorporated the marking key into the EPSS and developed and refined the views. Created two identical versions: a peer group EPSS and a tutor EPSS.</td>
</tr>
<tr>
<td>4: Pre-marking moderation</td>
<td>Coordinator</td>
<td>Selected a sample range of the students’ event analysis maps to be moderated during a regular unit team meeting before the tutors started marking.</td>
</tr>
<tr>
<td>5A: Peer group marking</td>
<td>Team</td>
<td>Decided how best to manage the student peer assessment process (e.g., removing cover sheet, coding papers, printing out blank rubric recording sheets).</td>
</tr>
<tr>
<td></td>
<td>Students</td>
<td>Each group member orally presented their event analysis map to the group, then left the room while the group first recorded their individual marks, and then moderated a group mark. This moderated group mark was at the end of the tutorial recorded via the Intranet into the EPSS.</td>
</tr>
<tr>
<td>5B: Marking</td>
<td>Researcher</td>
<td>Took the peer group EPSS and made a duplicate of it***.</td>
</tr>
<tr>
<td></td>
<td>Tutors</td>
<td>Marking occurred.</td>
</tr>
<tr>
<td>6: Post-marking moderation and re-submission</td>
<td>Researcher</td>
<td>Recombined the individual tutor EPSS into a unit EPSS.</td>
</tr>
<tr>
<td></td>
<td>Coordinator</td>
<td>Carried out quality assurance and control and returned student feedback sheets to tutors.</td>
</tr>
</tbody>
</table>

*Note: activities performed by students are shown in italics.*

At the moderation meeting, the decision was made to incorporate peer group marking into this DC. However, this time the peer group mark would not be part of the moderation process but would form part of the total student mark. The unit team decided that all marks allocated to the *Oral Presentation* (7 marks out of a total of 30) would be derived from the peer assessment, while the remaining 23 marks were allocated to the *Event Analysis Map* marked by the tutors.

Activity 3 incorporated the signed off marking key into the EPSS by the researcher. Due to the student participation in the task assessment process, the standard views developed in previous DCs were modified; these views were: *student feedback* view (Figure 7.7), *tutor marking* view, *spreadsheet* view and *peer group marking* view (Figure 7.8).

The tutors marked the *event analysis map* out of 23 and the student peer group marked the *oral presentation* out of 7. An example of a combined mark and grade is shown in Figure 7.7 (top right corner). The *tutor marking* view with a global comment box and the *spreadsheet* view were created with linking buttons to allow the tutors to
move between these two views just by clicking the button next to each student ID number. In the previous development cycle (DC4), only a one-way button had been used and based on the positive feedback from the tutors the additional button was created. The peer group marking view was used to record via the Intranet the student peer group oral presentation mark. This view also included the analysis map section that the peer group could also record their marks in, however these marks would not be used by the tutors for moderation purposes.
Activity 4, pre-marking moderation, involved the coordinator selecting a range of analysis maps and tutors using blank marking keys to mark them during a normal unit
team meeting. After marking, the tutors discussed and came to a consensus about the marks and relevant comments.

Activity 5 consisted of two parts: the student peer group marking and the tutors’ marking. Activity 5A, student peer group marking, took one tutorial session to complete. Five to six groups of five students were arranged around the room with desks. Students were each given five minutes to present their analysis map and then left the room while the other members of the group determined a group mark. When this was completed, they called the student in and the next student presented. At the completion of the students’ presentations, the moderated marking keys were entered by the students on the Intranet. Finally, the moderated marking keys were handed in with the Event Analysis Map.

**Figure 7.8 Peer group marking view (DC5).**

Activity 5B, tutor marking, firstly linked the peer group EPSS to the tutor marking EPSS. A copy of this combined EPSS was then copied to the tutor’s computer and the procedure for marking and navigation explained.

Activity 6, post-marking moderation, was again carried out by the coordinator with the assistance of the researcher. A number of summary views were developed by the researcher during this DC to help the coordinator carry out quality control of the student marks (two examples of these summary views are shown in Figure 7.9).
Discussion of findings from DC5

The combining of Activities 1 and 2 worked well. The process of using the two previously developed marking keys, the oral and poster presentation, to develop the working copy of the event analysis map by the coordinator was a success. The tutors found that the e-mail process and the moderation meeting that followed was efficient and produced a good marking key.

The researcher found Activity 3, the incorporation of the marking key and completion of the different view into the EPSS, complex. However, the researcher was able to build upon the experiences and knowledge gained from the previous DCs. The Intranet access for the recording of the student group marks (Activity 5A) again worked well with access to the five computers within the tutorial classroom greatly facilitating this process. The linking of the two EPSSs and combining of the student group marks and comments with that of the tutor’s mark was found to be seamless, but did require some thought for the researcher on how to link the two EPSSs.

Activity 4, the pre-marking moderation, was again a success with the use of the blank student feedback view being used as the recording sheet by the tutors. As the student groups and tutors marked different sections of this assignment there was no ability for the tutors to use the student group mark and comments for moderation as had been done in the previous DC.

The marking process (Activity 5) was again enthusiastically received by the tutors. They found the buttons linking the two views, the tutor marking view and the spreadsheet view, a “useful additional feature”. A problem that arose early in the marking for some tutors was that the students group mark represented “too great a
proportion” (seven out of thirty or 23% of the total mark) and in some cases distorted the total mark for the assignment. To help overcome this, a field was added to the tutor marking view that allowed the tutors the ability to add marks to a student’s total, up to a maximum of 2.5 in increments of 0.5. This was not a difficult adjustment to incorporate but required the adjustment of the five tutor EPSSs individually. However, in reviewing the marks, it was found that adjustments occurred in only 30 out of the 215 students.

The coordinator again found Activity 6, the post-marking moderation of quality assurance and control, an efficient process. The ability to view and spell-check all the tutors comments and sub-marks was again considered useful from both a student learning and teaching perspective.

**Written exam development cycle (DC6)**

**Description of assessment task**

A written examination was the final assessment in Semester 2, and it was worth 50% of the unit total mark. The sources of data for this development cycle are summarised in Appendix A. The two-hour written exam occurred during the two-week examination period at the end of the semester. The students were required to develop written responses to two questions from three published during the semester.

**Question one: 30 marks (this question is compulsory)**

A teacher can play a significant role in supporting children’s social and emotional development:

- Identify and describe a range of ways teachers can provide positive support.
- Illustrate your ideas with examples drawn from the K-Primary age range.
- Make explicit links to relevant theory.

**Question 2: 20 marks**

Describe the significant influences on a child’s development of a sense of self (self concept, self esteem, self efficacy):

- Discuss how this sense of self might affect the child’s educational outcomes.
- Make reference to appropriate theory in your answer.

**Question 3: 20 marks**

Describe the interrelationship between the cognitive, social and emotional dimensions of learning:

- Illustrate your ideas with examples drawn from the K-Primary age range.
- Make explicit links to relevant theory.
Description of activities

As the unit team had not previously developed an EPSS for an examination assessment task, they were unsure whether an EPSS would be of any benefit or advantage to this task assessment process. However, based on the team’s previous experiences of the benefits and advantages of using an EPSS, they decided to explore the possibilities.

A summary of the activities involved in this DC is shown in Table 7.5.

During the first unit team meeting (Activity 1), team members discussed how the assessment of examinations had been performed in the past. The assessment of examinations had usually involved the coordinator developing a rough marking guide that was supplied to the tutors with the students exam papers. Tutors had been required to elaborate on this limited marking scheme and develop their own procedure for recording the marks and returning them to the coordinator.

The unit team decided not to develop the usual detailed marking key. This was in part due to the limited time available to complete the DC, approximately two weeks, and the fact that no feedback or marks were returned to the students. This meant that a number of activities included in previous DCs were either eliminated or reduced in time and complexity. The unit team decided that to help reduce discrepancies and variability
between markers each exam question would be marked by only two tutors where possible.

At the meeting, the team also decided to eliminate Activities 2 and 4 from the DC. The unit team developed criteria headings including the general descriptors for each criterion for the compulsory essay question. The criteria headings and marks allocated were the following:

- **introduction (4 marks):**
  - brief outline of the scope of the essay, containing the main points that will be addressed, briefly introduce/define any key terms;

- **body (22 marks):**
  - identify (10 marks)
  - illustrate (6 marks)
  - make explicit links (6 marks)

- **conclusion (2 marks), and**

- **legibility (2 marks).**

Based on the criteria headings and general descriptors developed for the compulsory question, the two remaining essay examination questions marking keys were developed. The coordinator once more was responsible for collating and editing this material, and produced a working copy of three examination questions marking keys.

Activity 3 involved the researcher incorporating these three examination questions marking keys into the EPSS. The researcher also used the notes from the meeting to further refine the design of the EPSS to create the following views:

- **Q 1 marking view**
- **Q 2 marking view** (Figure 7.10)
- **Q 3 marking view**
- **Spreadsheet view**
- **Unit spreadsheet view** (Figure 7.11)

The researcher decided to incorporate this EPSS consisting of the three examination marking keys into the EPSS developed for DC4. A number of factors contributed to this design decision. One factor was the limited amount of time available to development a new EPSS. Another factor was that that these marking keys were relatively simple in design compared to the detailed marking keys developed in the
previous DCs and thus the three new marking keys could easily be incorporated into an existing EPSS.

Each of the essay examination marking keys had its own marking view that was similar in layout. In place of grade descriptors and grade word marks (e.g., ‘Pass’), coloured bars were used to approximate grade levels and numbers were used to record the mark. Navigation buttons, based on the favourable feedback from previous DCs where they were first introduced to a limited extent, were again used. The buttons were used to inter-link the three examination questions marking views (see *Q2 marking view* in Figure 7.10) and to link them to the spreadsheet view. The *unit spreadsheet view* (see Figure 7.11) was an expansion of the assessment spreadsheet view, and included all the assignment marks and the unit total mark and grade. At the completion of the development of the EPSS, this was duplicated five times and one copy was given to each tutor for marking.

During the marking activity (Activity 5), the coordinator divided the three examination paper questions equally among the tutors, based on their workload. Where more than one tutor marked the same examination question, the tutors carried out an internal moderation process, thus improving the reliability and consistency of their marking.

Activity 5 proceeded satisfactorily with all tutors using the EPSS to record the essay examination marks. Tutors collected their group of examination papers from the coordinator or collection point, soon after the completion of the examination. If tutors had not already collected the EPSS, they obtained their copy when they collected the examination papers. The turn-around time was less than two weeks and included the moderation of the examination marks, the combining of all the assignment marks and the submission of the total student unit mark to the Board of Examiners.
Chapter Seven: Development Phase - Semester Two

The post-marking moderation (Activity 6) was greatly reduced in this DC because no feedback to students was required. This meant that quality assurance was greatly reduced with quality control only involving checking that the spread of marks within and between tutors and overall, and also between examination questions, was acceptable. The need for quality control was further reduced by the limited number of tutors involved in marking each question and by their collaboration and moderation during the marking process. The EPSS made these comparison checks quick and easy due to the **find** and **sort** functions within different views.

As this was the last assessment item for the semester, the coordinator needed to combine this assessment examination mark with the other two assessment marks to produce the total unit mark for each student. The researcher was able to use the EPSS to link and add these three assessment marks to produce the unit total mark (see Figure 7.11). This view allowed the coordinator to see the three assessment marks, total mark, and grade for each student in a *spreadsheet* view. The coordinator was then able to quickly and efficiently carry out quality control of the unit mark including the unit overall spread of marks and the individual spread of assignment marks. Finally, the unit list was sorted by student surname and printed, and the student unit total marks from this printed list were entered into the official university unit list. This document was then processed as per the official university requirements.

---

**Figure 7.10 Exam Q2 marking view (DC6).**
Discussion of findings from DC6

Activities 1 and 2 were combined and occurred during the same team meeting. The moderation process involving the design and edit of the marking keys proceeded smoothly and efficiently. In part, this was because the team members were now used to working together, and they were familiar with the unit material and the examination questions. The coordinator, directly after the moderation meeting, collated and edited the examination marking keys that resulted from this moderation process. This document was then emailed directly to the researcher for incorporation into the EPSS (Activity 3).

Because these marking keys were less detailed than the ones developed previously, the development of the views was greatly simplified and the complexity of the EPSS was greatly reduced.

The tutors found that Activity 5, the marking of the student examination papers, went smoothly. The EPSS eliminated all the busy and unproductive work of adding the sub-marks and recording each student’s total examination marks.

The moderation process involved in Activity 6 went as planned, and the coordinator was pleased with the easy and efficient way they were able carry out the quality control and view the different student marks: the individual exam question mark, total exam mark, the other assessment marks, and the total unit marks and grades. This was in marked contrast to previous methods used to collate the three assessment marks and total them. The final step of copying the students total unit mark to the official spreadsheet document was however still done by hand. This last step could have to some extent be automated, but as it was the final process in the task assessment process it gave the coordinator time to reflect on the individual student marks ensuring that no errors had occurred.
Conclusion

The discussion of findings from DC6 completes the discussion of the Development Phase of the study: the iterative development of an EPSS prototype to support and enhance the task assessment process. The formative findings and results from the Development Phase discussed in Chapters 6 and 7 were encouraging and suggested that the EPSS and intervention strategies implemented enhanced the performance of the task assessment process. The Evaluation Phase contains the summative findings of the study. These findings are discussed in the following two chapters (Chapters 8 and 9). Chapter 8 discusses and evaluates the EPSS as a product, while Chapter 9 discusses the effect the EPSS had on the whole task assessment process.
CHAPTER EIGHT
EPSS EVALUATION

Introduction

The Evaluation Phase of the study combined: 1) the findings of the formative evaluation that occurred throughout the Development Phase, and 2) the summative findings that arose from the semi-structured interviews that occurred at the end of the Development Phase. The findings of the Evaluation Phase are discussed in this and the next Chapter (Chapter 9). The overall effect of the implementation of the EPSS and intervention strategies on the task assessment process are fully discussed in Chapter 9, while this chapter discusses the findings concerning the performance of the EPSS as a product, focusing on the human computer interaction (HCI) and how this interaction contributed to the task assessment performance gains.

Licklider (1960) described human computer interaction as symbiotic, referring to the fact that humans and computers “cooperate to obtain a goal because each component has unique abilities to bring to bear on a given task” (Badre, 2002, p. 3). In the task assessment environment, the assessor brings their professional judgement and knowledge of the content and curriculum, while the EPSS provides rapid and accurate data manipulation, storage, retrieval and management, and analysis and display ability. The assessor and EPSS thus augment, complement and enhance “each other’s capacities in performing complex, multifaceted tasks” (Badre, 2002, p. 3). To achieve this performance augmentation and enhancement of the task assessment process, a design for use or User-Centered Design (UCD) methodology was used throughout the Development Phase to produce a ‘useable’ product. Cato (2000) describes this process as “one which is conceived and produced to be easy to learn, easy to use and useful” (p. 5) and emphasises the need to design iteratively. He uses the following definition to illustrate this point:

Design is solving problems that cannot be formulated until they have been solved. The shaping of the answer is part of the question. (p. 2)

The summative data was gathered through semi-structured interviews conducted at the end of the Development Phase. Individual interviews were conducted with the team members who were involved in the development cycles in Semester 2. A semi-
structured interview technique was used both to obtain comparable responses and to allow a conversation to develop between the researcher and interviewees.

The evaluation of the EPSS as a cognitive support or performance aid to the task assessment is discussed in this chapter. The coordinator and tutors’ reactions and responses to these performance enhancements related to the application of the EPSS, are evaluated as assertions in the next chapter, Chapter 9. This division of the performance findings was deemed appropriate because the full benefit of the EPSS could only be achieved and evaluated when the synergy of these other performance enhancements and practices are combined and integrated in light of the users’ response.

In researching the academic literature on assessment, the researcher found that the application of criterion- or rubric-based designs to the task assessment process was growing, but it appeared to be mainly used and driven by the primary education sector. This literature also showed limited examples of the application of EPSS or Computer Aided Assessment (CAA) tools to the task assessment process where professional judgement was involved. The existing EPSS or CAA applications had not involved the whole task assessment process, that is, from the design of the marking key to the management of the students’ marks. The existing EPSS or CAA applications instead seemed to be limited to specific aspects of the task assessment process, such as the collation of marks or recording of multiple choice questions (MCQ) marks. Pearson Education Australia, a leading publisher of educational resources, has brought together a comprehensive list of e-assessment products at their Internet site (Pearson Education Australia, n.d.). This list, however, also only focuses on specific processes involved in task assessment, with the focus being mainly on automated objective based marking.

The final prototype of the EPSS completed at the end of the Development Phase (see Chapter 7), showed that all activities involved in subjective task assessment could be supported and enhanced. The final version of the EPSS showed varying degrees of integration across all the activities of the task assessment process, from Activity 1 to 7, with differing degrees of performance enhancement.

The summative evaluation findings of the EPSS are discussed and grouped under the following themes: design, usability and implementation. These areas were to varying degrees explored and investigated in the previous two chapters (Chapters 6 and 7) with the aim of improving and enhancing the performance of the task assessment process. These summative findings build on the formative results of the Exploration
Phase (discussed in Chapter 5) and the Development Phase (discussed in Chapters 6 and 7), that were used to inform, refine and develop the EPSS prototypes.

**Design themes**

The key design features of the EPSS involved the division of the task assessment process into discrete activities spread over approximately six weeks for a typical development cycle (DC), with about two weeks of overlap between DCs. At the completion of the Development Phase, all the activities showed varying degrees of integration and blending between the intervention strategies developed and the EPSS.

The least degree of integration occurred during the design and development, and quality assurance of the marking key (Activities 1 and 2). This involved the tutors’ collaboration, to varying degrees, in the first two weeks of the DC. The highest level of integration occurred in Activities 5, 6 and 7, where the detailed marking key was embedded into the EPSS and provided different electronic views and methods of recording marks and comments. The features and functions of the EPSS resulted from collaborative decisions that the unit team made during previous DCs and activities. The uniqueness and usability of the EPSS can be attributed to the collaborative nature of the development methodology that is the UCD or participatory design (PD), with the coordinator and tutors as collaborators.

It is important to note that during the interviews, while one process, strategy, feature or function was being discussed and analysed, often other strategies from a previous activity were contributing and aiding the performance of that process. For example, when Tutor B commented on the advantages of marking with the EPSS (Activity 5), her comments are a reflection on the success of Activities 1, 2 and 3:

> There were a lot of benefits really, consistency, I was a lot more consistent with my marking than I was previously ... it took away a lot of the subjectivity ... and easy to mark using the laptop. Oh, the other good thing was that (student) records were easy to access. ... If I needed to revisit a (student) record for a particular student, in the past I had to shuffle lots of paper, and keep things in folders, this is [now] so easy to find a [student] record. [Tutor B]

These features or functions of the EPSS were developed within each activity and worked together to form a unique blend and synergy within which the EPSS was applied to the task assessment process. This resulted in a design that was complex but simple to use. The tutors all found this to be the case, which Tutor E put succinctly as, “What I real liked about it is that it was simple”.
The coordinator and tutors were the driving force for all the key design decisions and features as part of the UCD methodology employed in the development process. Design features were developed for each activity and, where appropriate, they were incorporated into the EPSS during or at the end of each iterative DC. The researcher’s knowledge of interface design and programming was used to reduce the cognitive load involved in using the EPSS in each activity. This was achieved, wherever possible, by having the screen (computer) view reflect what had been developed on the paper copy, that is, *what you see is what you get*. When this rule was violated, the aim was to increase functionality and performance or to enhance the paper version, that is, *what you get is more than what you see*.

During each activity within the DC, a number of design decisions and/or features or interventions were made and incorporated into the activity and where appropriate into the EPSS. These were designed to improve the performance of that activity or of a later one, and thus the performance of the task assessment process as whole. Often these decisions or interventions resulted in the elimination or removal of an activity or process rather than the creation of an additional process. For example, the elimination of the clerical and administrative work associated with the task assessment process resulted in an improvement in performance of the users. A summary of the final design decisions grouped by activity is listed in Table 8.1.
Chapter Eight: EPSS evaluation

Table 8.1  
Summary of the key design decisions and features by activity

<table>
<thead>
<tr>
<th>Activity</th>
<th>Name of Activity</th>
<th>Key Design Features</th>
</tr>
</thead>
</table>
| 1        | Design and development of marking key    | – Start at least four weeks before marking  
|          |                                          | – Coordinator to develop a working copy of the marking key (ideally with tutors’ input)  
|          |                                          | – Criteria to be based on the desired learning outcomes  
|          |                                          | – Four levels of grade descriptors  
|          |                                          | – One tutor comment box  
|          |                                          | – One global tutor comment box  
|          |                                          | – Word grades for recording criteria marks  
|          |                                          | – Marks for generic skills  
|          |                                          | – Student involvement where possible  
| 2        | Quality assurance of marking key         | – Start at least four weeks before marking  
|          |                                          | – Tutors to be involved in the quality assurance and refinement of the marking key  
|          |                                          | – Student involvement where possible  
| 3        | Incorporation of marking key into the EPSS | – Student and tutorial details imported into the EPSS  
|          |                                          | – Main views created:  
|          |                                          |   – Tutor marking view  
|          |                                          |   – Student feedback view  
|          |                                          |   – Spreadsheet view and  
|          |                                          |   – Unit total marks view  
|          |                                          | – Other views included:  
|          |                                          |   – Spelling view  
|          |                                          |   – Summary views  
|          |                                          | – Automation of processes  
| 4        | Pre-marking moderation                    | – Hard copies of the student feedback view used to record the pre-marking moderation marking  
| 5        | Marking                                  | – Tutor supplied the EPSS  
| 6        | Post-marking moderation                   | – Tutor copies of the EPSS combined to facilitate the quality assurance and control process  
| 7        | Management                                | – Linking of the task assessment marks to provide a unit total mark and grade  

The features built into the EPSS evolved throughout the six DCs. The features were built upon the feedback and findings from the previous DCs. This process resulted in a set of features that a mature EPSS prototype might require in order to support and enhance the performance of the task assessment process depending on the type of assessment method used. These suggested features as they evolved and matured through the DCs are summarised in Table 8.2.
Many of these design features were designed to eliminate all or most of the work that did not require the tutors’ professional judgement or thoughtfulness. This resulted in the tutors being able to focus their attention on making professional judgements and not being distracted by the unproductive clerical and administrative work involved in the marking process. These features were designed to produce major performance gains in productivity, both in time saved and in the quality and performance of the tutors’ professional judgements. The team all commented positively on these design features and felt that their marking and that of the other team members were fairer, more reliable, valid and transparent (see Chapter 9).

### Suggested additional design features

Few suggestions or comments came out of the interviews regarding the need for additional design or layout features for the EPSS. This could be attributed to the iterative nature and number of the DCs that addressed most of the suggestions and comments of the team during the Development Phase. One suggestion was the need to deal with students’ assignments that are handed in late, and another comment related to a minor technical issue concerning the rounding of sub-marks and marks. Both these issues could be resolved easily through modification of the EPSS. Another issue raised...
was whether to show the criteria maximum marks or group marks on the student feedback view. Tutor A felt that they should be shown because:

*They want to see what sections are worth more than others ... the need to see that the summary is worth more than the other sections so they know where they should be putting their effort ... They need to be aware of the weight.* [Tutor A]

Another modification related to the peer group marking suggested by Tutor A, was to have a recording box per criteria below the grade descriptor. This box would provide an area for the student to record their reasoning, rationale and justification for allocating the grade so that during the group moderation discussion, they would have notes to call upon. Similarly, the group would record their reasoning, rationale and justification for allocating the moderated grades per criteria.

Suggestions were also made concerning the issue of modification and adjustment by the user of the content, design and operation of the EPSS. The Coordinator and Tutor A specifically commented that they would have liked to be able to modify the content and design of the EPSS.

The suggestions from the Coordinator focused on management issues and the enhancement of the student information data. She suggested, for example, the incorporation of student photographs and the elimination of the role that the researcher (programmer) played in the management of the EPSS. This covered the setting up, importing and exporting, and analysis of the data.

*Collation of unit statistics need to be more streamlined, bit more manageable ... Start up process, photograph, holding area for general student information.* [Coordinator]

Tutor A commented that they wanted to be able to change the mark allocation more easily:

*I could change the wording, but I don't know whether I'd be able to change the marks, the mark allocation, I don't think I'd be able to do that, you'd have to show me how to do that.* [Tutor A]

These suggestions could be incorporated easily into the EPSS, with many of the modifications, adjustments and controls brought up to the user level. For example, the criteria marks could be adjusted by the coordinator at the user level of the EPSS. These suggestions and comments are all concerned with what is known as the usability of the EPSS.
Usability themes

Usability as it relates to product research can be viewed from two perspectives. Firstly, usability is the study of the behaviour of users and the discovery of what works. This was achieved in this study by the application of PAR (participatory action research), UCD and the Human Performance Technology (HPT) model. Secondly, usability can be considered as the belief that humans should have mastery over the constructed environment. The researcher held this belief and this fact contributed to the increase in the usability of the final product.

The International Standards Organization (ISO) model ISO 9241 (Bevan, 2001) defines usability as measures of the following four constructs: efficiency, as the accuracy and completeness the user achieves with respect to the goals; effectiveness, as the user effort required to achieve the user and domain goal; satisfaction, as the measure of user satisfaction on a number of attributes; and usefulness, as the measure of the value the user places on the product. This ISO definition is operational in nature, requiring task definition, user definition, and the means for measuring effectiveness, efficiency, and satisfaction. This study did not permit the quantifiable measurement of these attributes of usability but did permit the qualitative measurement through the iterative DCs and the semi-structured interviews. Users experienced products holistically and thus the evaluation of usability must be applied to the whole product.

These usability attributes have been used to group and analyse the findings to highlight where the performance gains were achieved. In all these areas of performance gains, the collaborators rated them as favourable or very favourable throughout the interviews. The use of enthusiastic responses such as: “the beauty of it”, “it was really easy” and “oh absolutely, incredible, much quicker” seemed to indicate delight, which is more than mere agreement or acceptance of this innovative product. Delight is the combination of the right amounts of the cognitive, behavioural and affection that the user has towards a product, and was referred to previously in Chapter 4 (see Figure 4.2). These comments lent support to the choice of the UCD methodology that was used throughout the Development Phase.

Efficiency construct of usability

The interviewees all commented that they found the EPSS product increased their performance and productivity in completing the activities involved in the task assessment process. Specifically, they were able to improve the accuracy and
completeness of the activities involved in the DC. The EPSS and the processes implemented during the Development Phase presented many opportunities for performance improvement through both improved efficiency and by eliminating inefficiencies throughout the iterative Development Phase.

The spacing of the activities involved in the DC over about six weeks, not only meant that they were carried out more efficiently but also meant that the coordinator and tutors were under less stress when the time came for marking because they had already increased their knowledge and understanding of the marking key. This understanding reduced the cognitive load and processing that was usually associated with this highly stressful and time-demanding aspect of the task assessment process, the marking activity.

The automated processes incorporated into the EPSS removed much of the routine, busy and unproductive administrative and clerical work traditionally associated with moderation, marking, feedback, reporting and management processes. For example, the EPSS removed the need to manually add student and tutorial details to the marking sheet, spell-check the comments and add up the scores. This automation eliminated the possibility of human error when performing these activities, thus improving accuracy. All the interviewees commented that they especially liked not having to perform these tasks. Not only was human error reduced but also the time involved in these administrative and clerical work tasks was substantially reduced, as the following quotes from the tutors demonstrate:

*In the past, ... I was terrible at adding up, I double-checked everyone, I would have been nervous that I had made a mistake. [Tutor E]*

*It (EPSS) saved an enormous amount of time in marks allocation and adding up, which I really valued because that’s something I do very slowly. [Tutor F]*

Tutor F, while reflecting on the performance improvements that the EPSS made to the marking activity, elaborated in more detail about the educational significance and benefits that these performance gains allowed. She not only appreciated the time that the EPSS saved her in the marking process but also appreciated that this gave her more time to devote to the “important things about marking”, that is, to determine the quality and to think about what learning was being demonstrated by the student. These improvements also included higher quality student and teaching feedback, improved transparency and reliability, and more time for evaluating the student’s work.
The EPSS features and functions developed during the Development Phase of the study also improved the accuracy and completeness of the marking key, moderation, marking, feedback and management processes. The quality assurance of the marking key and its grade descriptors greatly enhanced the effectiveness of the marking key. All the tutors and the coordinator agreed that the time spent in Activities 1 and 2 at the beginning of the DC, that is, the time spent in the design and development of the marking key and its quality assurance, was productive. The team found that these activities improved the reliability and accuracy of the marking and saved a substantial amount of time by reducing the time spent during the marking and moderation of the final marks (Activities 5 and 6). The anxiety, stress and time off-task that they usually associated with the marking activity was significantly reduced or eliminated when they used the EPSS. Tutor E expressed this in the following words:

*At that point of marking which is usually a time-consuming process, saves a huge amount of time ... The actual marking process (the EPSS and rubric) makes it a lot quicker and easier, and a lot fairer for the students and yes, lots less moderation at the end needed.* [Tutor E]

**Effectiveness construct of usability**

The effectiveness of a product refers to the ability of the user to complete tasks with reduced effort. In the case of task assessment process, the effort was reduced by the use of the EPSS in several ways. One of the areas of greatest reduction of effort was the reduction or elimination of the clerical and administrative activities associated with the task assessment process. Not only did this feature improve accuracy, as shown in the previous section, but the design of the EPSS also meant that the effort involved in the process was greatly reduced.

The EPSS was able to simulate the paper-based versions of the student feedback sheet developed in Activities 1 and 2 and a typical spreadsheet layout. This almost identical appearance meant that the effort in terms of learnability and cognitive load was greatly reduced. This was reflected in the interviews with the issue of the need for training or problems of remembering how to use the EPSS not being raised. Once the process of opening the EPSS had been learnt (and this involved only clicking on an icon and entering a login and password), the tutors only needed to remember how to record the *grade-word* mark. This was via a mouse click and once shown, the operation became intuitive. Once the *grade-word* mark was recorded, the EPSS converted this into a numeric mark and added this to the assessment total seamlessly and
automatically. The recording of tutor comments and global comments were similarly intuitive. The navigation between the three main views of the data was via clicking on coloured buttons and was quickly learnt and remembered when the next assessment was marked approximately six weeks later.

The ability of the EPSS to automatically display different views of the data, such as the spreadsheet view, resulted in substantial savings in both effort and time. This and the ease of finding students in the different views also reduced the effort involved in the task assessment process. Tutor B put this in the following words:

*You can collate all the marks, and all that data stuff was done by the computer. ... The other good thing was that records were easy to access, if I needed to revisit a record of a particular student. In the past I had to shuffle lots of paper, and keep things in folders, this is so easy to find a record.* [Tutor B]

### Satisfaction construct of usability

Satisfaction of a product or process is a subjective evaluation by the users and refers to how pleased and satisfied they felt after using the product or process. In this study, the product was the EPSS. All team members in this study, coordinator and tutors, commented that, when using the EPSS and the processes and strategies implemented during the Development Phase, they all had highly satisfactory experiences. Coordinator and tutors attributed this satisfaction to the following:

- Improved confidence in making assessment judgements;
- Improved sense of control;
- Efficient use of time;
- Automated recording of marks;
- Improved management of the marking process;
- Eliminated unproductive work; and
- Improved modifiability of marks and comments.

All of the above areas are covered elsewhere in this chapter except for the area of improved professional confidence. This improved professional confidence referred to professional judgement, presentation of student feedback sheet and reduced stress. This was raised in all the interviews and was a cumulative response to this completely new approach of using an EPSS to aid and support the performance of the task assessment process. Tutor E commented on this improved confidence in terms of professional satisfaction about the marking activity and the results:
I don’t think I have ever marked in a fairer way, than doing it this way. I mean I think this is the fairest process I have ever been through in terms of marking with a group of people other than if you are just marking yourself. [Tutor E]

While Tutor B described her satisfaction with using the EPSS in these words:

It takes a lot of pressure off because at the end, you don’t have to go, “Oh is this a C or a D”, ... It's so easy, so much easier. And going back and changing marks, it's so easy to go back and just change marks. ...[Tutor B]

At the tutor level of the marking activity they all commented that they felt more in control of the marking activity when using the EPSS than in the past when using a paper based system. Tutor B put this in the following words:

You know, the beauty of it was really that at any stage I could go back and have a look at the overall marks and switch between views because we had to have so many As ... the grade distribution, I could be checking that as I'm going and just by clicking buttons, not having to do any adding, I wouldn't have access to that ... it’s quite like a luxury button. [Tutor B]

The coordinator and tutors felt that the EPSS supported, enhanced and improved their skills and expertise. The EPSS not only improved the efficiency and effectiveness but improved their overall satisfaction with the task assessment process, as the following two comments reflect:

I marked those assignments quicker than I have ever marked any assignments, because all the headwork had been done. I wasn’t doing it as I went along. And I felt far more secure about how I had marked them than other assignments. I think the process and instrument were fine and I think they worked really well. [Tutor F]

Oh, the spellchecker was good. My typing is not up to scratch so the spell check was really good ... Yeah, keeps it all uniform ... it shows that it's thought through, and beside that the students are getting feedback from the rubric itself, the whole thing is very well thought through and planned, and organised, and when you're organised it's easy to sit down and mark. If I feel organised, I can sit down and mark. [Tutor B]

**Usefulness construct of usability**

Usefulness is the value users place on a product or process. In this study, the product was the EPSS. The team all commented on the product’s usefulness to them in the task assessment process over the previously used manual methods. They were more efficient and effective in carrying out the activities involved in the task assessment process. The coordinator put this in the following words, focusing on how it benefited them in the real purpose of assessment:
Big benefit for me was that it reduced the busy work. So I was able to spend the
time actually considering the student work, instead of doing the busy work of
writing the name on the sheet, writing of the marks, writing the comments.
Because the bottom line is you have only x amount of time for each paper that
you are marking. So, if your time is taken up doing that busy work you have less
time to spend on actually making judgements about the students work. So, for
me the big difference was that I was able to spend more time doing what I think
assessment is about, feedback. [Coordinator]

All the tutors made similar comments to this in regards to the usefulness of the
EPSS in the task assessment process. They all raised and mentioned similar advantages
of using the EPSS over the previous manual methods. Tutor E put it like this:

The automatic adding up of the marks, the student names being there
automatically, just saves so much time. [Tutor E]

And Tutor A stated:

It was just a lot more user-friendly in that you didn't have lots of pieces of
paper lying around, it was just the assignments and the actual computer, the
laptop, that's all that was involved. [Tutor A]

Implementation themes

The implementation issues of the EPSS cover the observations made by the
researcher during the Development Phase and the findings based on the interview data
arising from the Evaluation Phase. The following themes emerged from the interview
analysis and are used to discuss the implementation issues:

• Attitude of participants;
• Training;
• ICT experience;
• File management, version control and maintenance;
• Distribution and collation of data; and
• Future directions

Attitude of participants

The researcher could not have asked for a more enthusiastic, cooperative and
collaborative group of tutors and coordinator to undertake this intensive one-year
participatory action research study. The Coordinator was always willing to dedicate the
extra time and effort to collaborate on exploring and developing the features of the
EPSS and DC activities. Over the Development Phase, the EPSS underwent many
changes and refinements, due to the active and collaborative involvement of the team.
Even during the most stressful times of marking, deadlines, and other outside commitments, the researcher felt that the team members were genuinely working as a team, and showed a high level of enthusiasm and commitment to the research study. A good example of this enthusiasm and willingness to try new features and possibilities, and overcome initial fears was Tutor B’s comment about first starting out in the project (Tutor B had only just started to use a computer at this stage):

Well, it was scary to start with. I lacked a lot of confidence to start with, because I really didn't know. Once I worked out how it works, I saw the advantages of it, but to start with, it was a bit daunting, way back at the start of the year. [Tutor B]

Tutor B reflected on the unit team’s collaboration and support for each other in the face of their lack of ICT skills:

There was a lot of support and that is the thing that got me through. I don't think I would have ... the first instance, I thought that everybody knew more than me and then when I realised that Tutor A doesn't, I felt kind of relaxed, and then I realised that because this is new to everybody. ... I still thought about learning, more about computers than I did. ... Struggling and I'd never used a Mac before, ... but I felt that I had the support and I could ask silly questions if I had to. ... And sometimes, when I asked silly questions I found that other people had the same questions, so I didn't feel so bad, I felt confident to do it. [Tutor B]

When the interviewees were questioned about the iterative and collaborative nature of the development process, the overall feeling was that because the functions and features evolved slowly and under their direct control, they were more able and willing to use them. When discussing the addition of new features to the EPSS, Tutor B made the following comment:

... Because we did it that way, I was more likely to use the flashy stuff. Had it all been presented to me at the start, there's a possibly that I wouldn't have used them because I would have thought it's too much and I wouldn't have gone there. ... You scaffolded my learning. [Tutor B]

Training

When commenting about professional development or training in use of the EPSS the team members agreed that due to both the iterative nature of the development process and having the researcher and other team members on-call eliminated the need for formal training during the Development Phase. Tutor B responded in the following words when asked about first using the EPSS and the question of whether she would have liked a workshop on the use of the EPSS:
Well, to do that I went home ... I clicked a few buttons and when I mucked it up, I went to see you and I asked you why is this not working and you showed me. ... No, it wasn't necessary in that case. Whenever I needed to, I came and asked. And what I didn't know, I worked out for myself. .... No, we never fell behind. [Tutor B]

In the post-Development Phase, training in the use of the EPSS was also found to be minimal, as the computer screens simulated the paper copies very closely and this reduced the cognitive load and the need for training.

**ICT experience**

The tutors’ inexperience of ICT applications did not substantially affect their performance gains during the task assessment process when using the EPSS. The team members’ prior experiences of using ICT were useful in interpreting their responses to the use of the EPSS. One might expect inexperienced ICT users to be more demanding and critical of the usability and less forgiving than more experienced ICT users. As it happened, the researcher’s observations throughout the Development Phase and the interview findings revealed that the team members were all either novice users or advanced beginners in ICT usage at the start of the research study. This was supported by the findings from the Exploration Phase (Chapter 5), which indicated that most tutors, lecturers and coordinators within the School of Education were novice users of ICT. However, the researcher observed that all the team members improved their ICT skills as the development cycles proceeded.

The interviewees’ responses concerning how they had marked in the past indicated that their previous marking activities involved very little use of ICT. Most tutors tended to use a Microsoft Word template of the marking key that was given to them by the coordinator in either hard copy or in an electronic form. The required number was either photocopied or printed, and then the students names, the sub-marks and comments were recorded by hand and added up manually to obtain the total mark. These were the types of activities that the EPSS was designed to eliminate. The exceptions were the coordinator and Tutor F, who had been exploring the use of Microsoft Word to record students’ marks and comments electronically. On reflection, the coordinator was bemused with her early attempts at using ICT in the assessment process as the following quote highlights:

*I did get to a stage where I did use a word processor to record comments, so I was not writing the same comments, I was cutting and pasting, typed a mark in,*
printed it out and then added up by hand, multiple documents [laugh], no real direction or support in how to do this process. [Coordinator]

Although final unit marks had to be submitted on a spreadsheet, none of the tutors used the functionality of the spreadsheet. All tutors, by typing or by hand, manually entered the marks into either a Microsoft Word table or Excel spreadsheet for each assignment and the total unit mark was similarly calculated manually. The two main reasons for their lack of sophistication in using these applications seemed to be that they were novice users of all forms of ICT and they had not been given any support or guidance in when and how to use these and other applications. For example, they did not realise that Excel spreadsheets can be used to automatically calculate groups of numbers to produce totals and other mathematical calculations. Tutor F had been exploring the use of Microsoft Word over the last six years to record marks and comments electronically, but she had yet to progress to using mail merge and therefore manually typed in the students’ names and added the marks manually.

The lack of ICT skills and experience of the team members was not a major issue during this study. For the whole year of the Development Phase, there was a good working relationship built between the researcher and the team members, perhaps because they were collaborators and were involved in the major design aspects of the EPSS. Their involvement in the study may have increased their ICT skills and confidence as illustrated by the following comment by Tutor A about her own and the students’ responses to using the EPSS:

It gave me a lot more confidence with ICT, it’s been really good. ... I think that was something, even in the classroom where students were very anti-IT and not confident when you said, OK well go on, do it on the computer [peer group marking], because it was so easy, it gave them confidence to go in the computer room and use it. I think that was very beneficial for the students who are very hesitant about using computers. A lot of them really gained confidence just by doing that... [Tutor A]

The only obstacle mentioned by Tutor A was her lack of access to a computer because she had to book a university laptop for the marking activity. Tutor A explains how this may have reduced the efficiency in using the EPSS:

Well the only real obstacle I had was not having my own laptop because it was so convenient having, using a laptop. I had to book it out and that sort of thing ... But other than that, I can’t see any obstacles. ... I’d prefer to use my own laptop because I could either, be working at uni or at home in my own time. You set it up really quickly, you do a little bit here and there. The other sort of marking you really had to set aside a time and go and do it, whereas this way...
you could actually set it up and just mark a couple and close it down, it wasn't a big deal, it was easy to do just little bits. [Tutor A]

As an aside, Tutor A has now purchased a laptop based largely on the positive experience of using the EPSS.

In terms of personal development, all tutors felt they had gained in ICT skills and confidence. This was an important but unplanned secondary effect of the study that, as Tutor B explains, was not necessarily easy:

Once I overcame my initial fear, I think sometimes it was still a little bit overwhelming because I couldn't quite understand what you were doing with the tool. You know you made me sit there and watch you, and even though I didn't understand, I was learning something about the actual program. ... I'm more willing and I'm not as ... you know that initial fear? I haven't got that as much, I'm more confident to approach this [ICT]. ... Yeah. Oh yeah definitely. That's a bonus yeah. It's given me the confidence to say, “Ok I can do this, with other things”. [Tutor B]

File management, version control and maintenance

File management and version control was mainly an issue for the coordinator and researcher as they developed different versions and managed the tutors’ unmarked and marked copies of the EPSS. Although this was not difficult and was part of the process of prototype development, it did require thought and consistent use, as both the researcher and coordinator found. Often ad hoc naming conventions and version control were used due to tight development and time constraints that the coordinator and researcher were placed under. Over time, this became more difficult especially when different versions of the marking key were being developed and tested. The coordinator commented that:

Keeping track of which is the latest version, and where do you store them and version control is a major issue. [Coordinator]

The researcher, in discussion with the coordinator, developed naming conventions to cover the following different versions that needed to be created, distributed, collected and combined during the task assessment process:
• Blank EPSS (no student information) used to design the marking key;
• Unmarked EPSS containing all students (master);
• Individual tutor EPSS (unmarked);
• Individual tutor EPSS (marked); and
• Marked EPSS (master) containing all the students

This was not a major concern as the EPSS was designed to be used with only one unit; however, if this product were to be used across a number of units or a whole programme of study, maintenance and version control would need to be addressed.

**Distribution and collation of data**

Another important issue also related to file management was the two-way movement of recorded information between the coordinator’s master copy and tutors’ copies. This process was done initially with the assistance of the researcher. However, this process would need to be simplified, as this task must be accomplished by the users if the product is to be cost-effective. However, this issue was reduced and simplified throughout the **Development Phase**. Part of the solution was the use of a runtime version of the EPSS, which eliminated the cost of installing a copy of FileMaker on the tutor’s computer. These runtime copies of the EPSS could be transported and even operated off USB memory sticks. Currently the preferred process is for the tutors to email the completed EPSS files (they are usually 1MB in size) to the coordinator who then imports them into the master copy for that assessment. This process still involves a number of manual steps that could be automated in the future.

**Future directions**

Even at the completion of the study, both the coordinator and tutors were still imagining and exploring additional features and applications of the product. These investigations would not have been possible prior to the development of the EPSS. For example, the Coordinator mentioned that,

*One of the things I want to try is the tracking the students, not only in my unit. I want to see over time whether there’re using the feedback that we have given them, whether they are developing their writing skills, for example. Somehow, it would be great to use the tool to keep track of that from assignment to assignment, and unit to unit. I can do it between Learning and Development I and II.* [Coordinator]
Tutor F described how she had tried to implement what the coordinator was suggesting but had found that the increasing class sizes, from less than 30 to over 200 was making the task laborious and difficult to manage and continue. She described in some detail what she would like the new product to be able to accomplish. She would like the linking of the feedback from one assessment task to the next, with the tutor being able to view these comments as they write the current feedback comment - thus a true implementation of feedback for both tutor and student. This linking would continue over semesters and years, as she explains.

*That would save me so much time and it means you have a developmental approach to your students that you give them the message that each assignment is not just this isolated thing. It’s evidence of their learning and if they’re not providing evidence that they have learnt, then you as a teacher would be very concerned, just like they would be about their students, about that lack of learning.* [Tutor F]

**Summary**

This chapter presented an evaluation of the operation of the prototype EPSS, which was developed to improve and enhance the performance of the task assessment process. The evaluation of the EPSS considered design, usability and implementation issues with the focus on the usability of the EPSS prototype. Improved performance of the task assessment process was achieved to varying degrees within all activities due to the application of the EPSS. The improved performance overall was due to the synergistic combining or blending of the EPSS tool and the innovative processes developed. The following chapter (Chapter 9) continues the discussion of these performance improvements with the focus on the innovative processes with the focus on the research question and subsidiary questions. These questions evolved throughout the study, as is consistent with the nature of collaborative and participatory action research (PAR), and with a product participatory design (PD). Based on these research questions, assertions are developed and discussed.
CHAPTER NINE
FINDINGS AND DISCUSSION

Introduction

The previous chapter (Chapter 8) focused on the evaluation of the EPSS itself in its role of enhancing the performance of the task assessment. This chapter continues this evaluation but takes the broad perspective of ‘what happened’ overall, from the design, development and moderation of the marking key (quality assurance of the marking key), to the pre-marking moderation, marking, post-marking moderation, and management processes through to the final student unit mark. That is, this chapter considers the additional effects of the EPSS and intervention strategies on the performance gap observed in the task assessment process, as identified and discussed in the Exploration Phase (Chapter 5).

This chapter begins with a brief review of the task assessment and the activities involved in this process designed to help the reader interpret the findings. These findings are presented as four assertions that emerged from the data collected through semi-structured interviews conducted at the completion of the Development Phase, and from the evaluation of the formative findings from the Exploration and Development Phases.

Task assessment

The task assessment was the process that was under investigation in this study. The performance improvement of this process at the workplace level was the goal of the study. The study used a modified and adapted HPT model as the framework to investigate and evaluate these performance interventions (see Chapter 3). Based on this HPT model, the study was divided into three phases: Exploration, Development and Evaluation Phases. The findings from the Exploration Phase of the study allowed the task assessment process to be divided into seven major activities:
Activity 1 – Marking key design and development;
Activity 2 – Marking key quality assurance;
Activity 3 – Incorporation of the marking key into the EPSS;
Activity 4 – Pre-marking moderation;
Activity 5 – Marking;
Activity 6 – Post-marking moderation; and
Activity 7 – Management.

These activities were usually spread out over approximately six weeks (Figure 9.1), and they were used to give structure to the task assessment process, help inform the study and analyse the data.

These key activities were the areas where intervention strategies were considered and selected for implementation to improve and enhance the performance of the task assessment process. This consideration and selection began with the identification of possible intervention strategies in the Exploration Phase that were then incorporated to varying degrees and integrated into the EPSS developed during the iterative development cycles of the Development Phase, as described and discussed in Chapters 6 and 7. The aim of the study was twofold: firstly, to develop intervention processes and strategies within each of these activities that would enhance and improve the
performance of the task assessment process; secondly, wherever possible, to further improve the performance through the application and integration of these into an EPSS. During each development cycle the goal was to provide a seamless, transparent and comprehensive EPSS to enhance and support the performance of the task assessment process.

The conceptualisation of this combination and blending of the intervention strategies with the EPSS could be viewed as two layers sitting below and supporting the task assessment activities (refer to Figure 9.1). The goal was to have the EPSS combine and blend in with each of these activities in a seamless and transparent way in order to improve, enhance and augment the performance of each of them individually and the performance of the task assessment process as a whole. Often the improvements, enhancements and augmentations were achieved in a later activity, for example, the performance results of the developed marking key where only realised during the marking activity.

The performance enhancements of the different activities associated with the task assessment process often changed as the assessment method changed. These activity changes brought about both unique and generic changes to both the processes developed and the features and use of the EPSS during the Development Phase, as discussed in Chapters 6 and 7. The final version of the EPSS showed the highest degree of blending and integration, especially in Activities 4 to 7.

Some of the interventions developed and implemented within and during each of the seven activities were at times difficult to separate from the EPSS (see Chapter 8), that is, some of the intervention strategies were fully reliant on the EPSS and could not have been implemented independently. They were fully embedded and integrated within the EPSS, for example, the electronic marking and reporting activities.

**Assertions that emerged from the findings**

The Evaluation Phase involved the individual team members (tutors and coordinator) being interviewed for approximately one hour each at the completion of the Development Phase. The interview method was a semi-structured approach (see Chapter 4), and the questions covered the marking process, the moderation process and the development of the marking key (see Appendix K). At the start of the interview, a copy of the areas to be covered, and a copy of the flowchart showing the seven activities involved in each development cycle (see Figure 9.1) were used to help focus
Four assertions were developed based on the evidence that emerged from all the phases of the study (Exploration, Development and Summative). This evidence was derived from the evaluation of the formative findings from the Exploration and Development Phases and the summative findings from the Evaluation Phase. These assertions added to the knowledge base in the area of task assessment. The first two assertions relate to the improved performance of the marking key as a performance aid, while the other two assertions relate to the improved performance of the worker, in this case the coordinator and tutor. The four assertions were:

Assertion 1: The marking key was improved and enhanced.
Assertion 2: The educative value of the marking key was improved and enhanced.
Assertion 3: The coordinator’s performance was improved and enhanced.
Assertion 4: The tutors’ performance was improved and enhanced.

These major assertions are discussed individually in this section with contributing evidence derived mainly from the interviews (Evaluation Phase), and to a lesser extent from the Exploration and Development Phases. Edith Cowan University’s definitions of roles and responsibilities of coordinator, lecturer and tutor will be used when describing performance enhancements to the task assessment process (the roles and responsibilities of lecturers and tutors are the same according to these definitions).

**Assertion 1: The marking key was improved and enhanced.**

This assertion looks at the marking key as a performance aid from the perspective of the assessor and leads directly to the next assertion, which deals with the educative value of the marking key from the student’s perspective. The marking key that resulted from the Development Phase was a cognitive or performance aid for the assessor that incorporated features of knowledge management. The knowledge of the coordinator, usually in collaboration with the tutors, of the desired learning objectives and the desired performance levels was captured and made transparent through the design and development of the marking key (an instructional rubric). The performance support offered by the marking key improved and enhanced the assessor’s performance of the task assessment process through its increased ability to capture and make transparent both the desired learning objectives and the desired performance levels within each
criterion of the assessment task. The following components were identified as contributing to and supporting this assertion that the marking key was improved and enhanced.

1.a The application of an instructional rubric design to the marking key for the task assessment.
1.b The *constructive alignment* of the learning objectives to the assessment task.
1.c Tutors’ enhanced ownership and understanding of the task assessment.

The performance of the marking key was improved by the application of an instructional rubric design to the marking key in Activities 1 and 2. This performance gain was in part due to the fact that the marking key rubrics that were developed were instructional and not purely a scoring marking key. Instructional rubrics are often co-created with students’ and/or tutors’ involvement, handed out before the assignment task is completed, used to facilitate peer and self-assessment, and used by tutors to provide feedback (see section on rubric in Chapter 2). All these aspects, features and uses of an instructional rubric were explored and developed during the development cycles.

The use of instructional rubrics greatly improved the performance of the tutors in the marking activity (Activity 5). The team members all commented that they marked more consistently (intra-marker reliability), and that this consistency was found between the markers (inter-marker reliability). Tutor B makes this point, with the last comment referring to both the marking key and the EPSS:

*There were a lot of benefits really, consistency, I was a lot more consistent with my marking than I was previously, because it [the rubric] forced me to ... it took away a lot of the subjectivity ... What I real liked about it is that it was simple [easy]. [Tutor B]*

Tutor A commented that the use of this type of marking key reduced the indecision and time spent trying to work out an appropriate mark:

*... It [the marking key] seemed to eliminate a lot of the doubt. You didn't have to procrastinate as much because it was clear-cut. [Tutor A]*

An interesting general comment to come out of this area of discussion from all the interviewees was that fewer students, if any, queried their marks in this unit compared
with what was usually the case. In the past students would often argue about the sub-
marks and over single or half marks when this type of marking key was not used. Tutors
attributed this lack of confrontation and disagreement with the students to the detailed
textual description offered by the marking key rubric that describes why students got
the sub-grade word marks. Tutor A added to this point commenting on previous
marking experiences:

Oh, yes, because it [previous marking method] was not clear, “Why did you
give me a high mark here, and here a low mark” … but the feedback [now] is
much more specific. [Tutor A]

1.b The constructive alignment of the learning objectives to the assessment task.

The performance of the marking key was improved and enhanced by the
constructive alignment of the desired learning outcomes of the unit to the assessment
task via the marking key. This was achieved through the design, development and
quality assurance of the marking key rubric (Activities 1 and 2). Thus, both the marking
key and the assessment task were aligned. During the development cycles, a number of
strategies and different combinations of contributors (coordinator, tutors and students)
were explored. This ranged from the contributors carrying out separate activities to
completely combining and integrating activities. The contribution of the collaborators
also ranged from only the coordinator to groups of students within tutorial groups. All
these intervention strategies were found to improve and enhance the design and
development of the marking key (Activities 1 and 2) and thus the performance of the
task assessment. The performance gains due to the constructive alignment cascaded
through all the proceeding activities and directly contributed to the improved
performance of the task assessment process. All the interviewees referred to the
significance of these two activities for the overall performance improvement of the task
assessment process. None of them had previously focused or spent this much time and
effort on the actual development or quality assurance of a marking key. Commenting on
this aspect, the Coordinator noted:

Very little discussion about what the criteria actually meant, limited
moderation, … sometimes using very global judgements, like outstanding,
satisfactory, not satisfactory, right down to using very specific marks but no
discussion and no descriptors. [Coordinator]

The common feature throughout all these intervention strategies was the
application of an instructional rubric (Assertion 1a) and the constructive alignment to
the learning objectives (Assertion 1b) through the design and the quality assurance of the resulting marking key (Activities 1 and 2). The quality assurance processes involved both the coordinator and tutors ensuring that the developed instructional rubric, both the criteria headings and grade descriptors, were representative of and aligned with the learning outcomes.

Tutor F, in discussing these interventions and activities, gave the following reasons for the success of the instructional rubric: the development of a common language, and the rubric giving a neutral and safe area from which to collaborate and develop a common understanding within the group. Tutor F expressed this in the following words:

*It helped us develop as a team a language in common because previously sometimes it's really difficult in marking to share things in a way that's accessible to inexperienced tutors. ... What the rubric did was give us somewhere neutral to put stuff. We could focus on refining that ... it depersonalised it in a sense and allowed people to share equally regardless of their experience, inexperience or status in the unit. And that's a really good process.* [Tutor F]

Finally, the time dedicated to these two activities (Activities 1 and 2) although substantial, was considered in all the interviews to be essential to the success of the task assessment process. Tutor E put it like this:

*More time in developing it (marking key), but in the actual marking process makes it a lot quicker and easier and a lot fairer for the students and yes, lots less moderation at the end needed.* [Tutor E]

Tutor A commented on the benefits and importance of doing these activities over a number of assessment cycles and the professional learning that occurred due to these activities:

*The very first rubric we set up, once we started using it, we saw how it wasn't as detailed as it should be, and by the time we got to the sixth one, we really had made it much more detailed. So certainly, we learnt as we went along. The more specific we were, the easier it was to use.* [Tutor A]

1.c Tutors’ enhanced ownership and understanding of the task assessment.

The task assessment performance was improved by tutors’ enhanced ownership and understanding of the marking key. This understanding included the content of the marking key and what was being assessed in the assessment task. This additional benefit came out of the design and development, and quality assurance of the marking key (Activities 1 and 2). The collaboration and sharing of expertise between the
coordinator and tutors during the quality assurance process of the marking key rubric contributed to the performance improvement of the later activities of the task assessment process. The tutors in the interviews attributed this to the improved ownership and understanding of the assessment task that developed during Activities 1 and 2. Tutor F elaborated on the benefit of this ownership and understanding:

*We could focus on that, refining that [marking key], it depersonalised it in a sense and allowed people to share equally regardless of their experience, inexperience or status in the unit and that was a really good process.* [Tutor F]

In addition, to the question of the benefits of using the rubric marking key and quality assurance processes, Tutor F continued the reflection on ownership and understanding, commenting that:

*It really assisted in getting uniformity across the group and it means the moderation process is much easier. It forces you to think really, really carefully, about what you are valuing in the assignment. ... Because you are actually assigning real marks, you are not making very general statements and saying look, the introduction is worth 3. You are being very very specific about which aspects would you give one mark to, which aspects would you give two, which aspects you would give three. So, you are actually dividing it up and thinking about it really critically.* [Tutor F]

The team members all commented that doing the quality assurance activity of the marking key was rewarding, beneficial and educative even though this process was time-consuming and new to them. The tutors commented that they were able to take ownership of the marking key and understood what was going to be assessed and how after performing this quality assurance. The team further commented that the time spent on refining and moderating the detailed marking key rubric was valuable and educative (Activities 1 and 2). This contributed to the effective and efficient performance of the marking activity (Activity 5), and more than compensated for the time spent in the development and quality assurance of the marking key. They also found that as they accumulated experience in the processes involved in the development and quality assurance of the marking rubric (Activities 1 and 2) and even though the assessment items were different, the time spent in these activities was reduced with each successive development cycle.

Tutor F, raised an interesting question that could not be answered in this study, whether the marking key (instrument) would be transparent enough to actually enhance the process for the assessors if Activities 1 and 2 were not carried out.
You wonder always with these things, is the learning in the doing or does it preside in the instrument [marking key]. So can we gain benefit from using a refined instrument or are we forced to always go through the process because of internal learning that happens. I think that’s always an issue with things that are seen to actually assist processes that are very complex. Is the assistance in the process or is it actually residing in the instrument to make that process easier and I think that it is always a dilemma. [Tutor F]

While Tutor B commented on the potential issue of lack of ownership and how it might affect the efficiency of the marking process in these words:

*Now, that was interesting when they [the students] did it because ... OK, when we wrote, when we had to say in the actual writing of the rubric, we had ownership and I felt that when the students picked up our rubric and read it, because they didn't have ownership, I didn't think that they comprehended it to 100%. And I wonder how the new tutors this semester will pick up the rubric, because they don't have ownership. [Tutor B]*

**Assertion 2: The educative value of the marking key was improved and enhanced.**

The educative value of the marking key builds upon the previous assertion, that the performance of the marking key was improved and enhanced. Assertion 2 focuses on the student perspective of the marking key, that is, their experience of the marking key and its educative value. This refers specifically to how the marking key can be viewed as a teaching and learning performance aid and can supports students in their learning process.

The marking key was a detailed instructional rubric that provided transparent, timely, detailed and constructive information about what the different criteria performance levels were. This detailed and specific information about what was being assessed and the different levels of desired performance outcomes encouraged and promoted independent learning and enabled the feedback and feedforward to be constructive, consistent and transparent for all learners. The analysis of the interview findings indicated that the educative value of the marking key was improved and enhanced. The coordinator put it in the following words:

*The development of the rubric helps not only the student understand what’s being expected of them but also staff in terms of what they will be looking for.*

[Coordinator]

The following components were identified that contributed to and enhanced this educative performance improvement of the marking key by the application of an EPSS:
2.a The use of word grades.

The educative value of the marking key was improved and enhanced by the use of word grades instead of mark grades within each criterion. This was made possible through the application of the EPSS, as the equivalent process using a paper-based method would have been prohibitively expensive and time consuming. The analysis of the interview data indicated that use of word grades instead of mark grades within each criterion improved and enhanced the educative value of the marking key, focusing the students’ attention on the performance level within each criterion. The elimination of the sub-marks under the grade descriptors and their replacement with word grades resulted from the quality assurance process (Activity 2), during the Development Phase.

The team felt this feature aided in the accuracy and educative value of the marking in a number of ways. The feature removed the focus from the sub-mark to the four grade descriptions of each criterion. The grade word was calculated to be the mid-point of that grade based on the criteria mark, which, they felt, also gave a truer sub-mark. The use of the mid-point grade mark also helped to eliminate final marks close to a grade cut off point, and this helped the coordinator during the final quality control of the marks.

In previous units, the marking key used a number range or, if grade descriptors were used, the possible numbers were not aligned with the university grades and the levels were not described. For example, a criteria mark out of 7 would produce a possible range of marks between 0 to 7, with a pass being between 3.5 to 4.5 (assuming that half marks were being accepted by the coordinator). Now, by using word grades and the mid-point of the grade being the mark, this would produce a pass being equal to 3.85 or 55% of 7, although this number would be difficult for a tutor to add up when they have five criteria and 30 or more assignments to mark. The selection of the word Pass on the computer screen was all that was required of the tutor for that mark to be recorded and added to the total.
The team felt this feature helped both the students and tutors to focus on the criteria and grade descriptors and not be distracted by the numeric sub-marks. This feature, although possible to achieve using paper or hard copy, would be too labour-intensive to be feasible, while the EPSS accomplished this seamlessly.

Tutor E explained the advantages of using word grades:

\[\text{Not showing the marks was a good idea. ... I agree with that comment [grade descriptor] and that happens to be a credit. ... That gets them away from looking at marks as opposed to criteria. I think that was a big advantage of it, it did focus on the criteria more so that the mark and with the criteria that also helped them with the understanding and ... less people quibbling about their marks. [Tutor E]}\]

While Tutor F commented on the increase in objectivity in the allocation of marks achieved through the use of word grades:

\[\text{The allocation of marks was very very precise and uniform. And so, it tied down the moderation process and stopped it kind of getting too subjective in a sense. Because you actually had to commit to changing marks if you wanted to change the weighting of something. It could not be like no this assignment definitely is not a credit, I know it's a distinction you had to actually say why you knew it was a distinction and what criteria you would give extra weighing to make it. [Tutor F]}\]

When the coordinator was asked about the use of word grades instead of sub-marks, she responded:

\[\text{I think it was educative for all of us working out what the criteria was worth, and then working out the difference between a pass and a credit, you know, what that was worth. And we accepted that a pass was the mid-point of the university grade pass. [Coordinator]}\]

2.b The use of generic marks.

The educative value of the marking key was improved and enhanced by the use of generic marks. This feature of the marking key came out of the quality assurance process during the early stages of the Development Phase. It added a range of marks from negative to positive sub-marks for generic and objective-based areas of the assignment, such as spelling, grammar, word length and referencing, which are expected of student academic writing at university. This feature was added to the marking key rubric in later development cycles. The benefits of this added feature to the marking key was that it improved transparency and consistency of the generic marks that are often hidden and interpreted differently by assessors. Thus, tutors and students
now knew what was expected and how many marks were to be deducted or gained for these areas of the assignment.

The team felt that it helped with consistency in marking, standards and showed the students what was involved in, for example, academic professional writing. Tutor A, commenting on this feature, felt that the use of this feature was beneficial for allocating marks towards these generic and objective areas of the assessment:

> It’s assumed there are certain academic requirements, which they are all aware of, as long as we go through that at the beginning, and then I think when they hand in their assignments, if those requirements aren’t met, they should lose marks. ... I think it made us raise the standards a bit. ... It makes them work harder on those little things, ... and perhaps they might be motivated to go and actually look and learn how to reference things properly, or use the spell-check or something like that. I think it was raising the standards. [Tutor A]

The use and innovative application of the EPSS in this study made this feature easier to implement and evaluate than would have been possible with a paper-based system. The cognitive load placed on the marker to incorporate these generic marks, i.e. having to deduct marks, without the aid of the EPSS, would usually prevent this feature being implemented.

### 2.c The feedback to students.

The educative value of the marking key was improved and enhanced through the increased quality of the feedback to students. The instructional rubric design of the marking key provided built-in feedback through the criteria grade descriptors of the performance level required, from pass to high distinction, for each criterion. This feedback was aligned to the desired learning outcomes, and was designed and developed in collaboration with the coordinator and tutors (Activities 1 and 2) at least three to four weeks before the marking started.

This feedback provided to the students through the marking key was educative in three ways. Firstly, the marking key was provided to the students prior to them completing the assignment. This gave students a transparent and fair means to guide and explain the aims and goals of the assignment and the performance levels the assignment required of them. This also enabled students to complete the assignment in a constructive and thoughtful manner and helped them to structure and allocate time productively towards the desired learning outcomes. The marking key gave them and the tutors a clear textual focus for an educative discussion to occur about the assignment goals and objectives. Secondly, the educative value of the *marked* marking key returned
to the student resided in the fact that it clearly showed the level or grade at which the student was performing, that is, the feedback was clear and meaningful. Thirdly, the educative value of the feedback was enhanced due to the design of the marking key that clearly indicated what the student needed to do to improve their performance to achieve the next performance or grade level. This is often called in the literature feed-forward.

2.d Peer group marking.

The educative value of the marking key was improved and enhanced through the peer group marking activity. The pre-marking moderation usually involved the tutors using a hard copy of the marking key and marking a number of samples of student work, which were then moderated. A number of different processes were explored, in part due to the different nature of the assignments during the Development Phase. One that proved educative for the students was peer group marking. Students recorded the peer group mark online and this was printed below that of the tutor’s on the feedback sheet. The EPSS did this seamlessly and involved very little additional work for the team or researcher. The results of the student survey (discussed in Chapter 7), and the comments from coordinator and tutors all indicated that they found the peer group marking activity of educative value. When asked about student peer marking activity, the Coordinator saw a number of advantages:

That it gives the students not only a taste of what it’s like to do assessment but it also does it using the technology for their future teaching and makes the link that technology can help you do this more efficiently. [Coordinator]

Tutor E commented on the value of the student peer group marking in the following words:

A sense of what they thought, it was amazing really, and I think it’s a testament again to the detail and understanding of the rubric, but we found that basically that the change between ours and their marks was not significant. They were pretty spot on generally. [Tutor E]

While a typical student response was very positive about the peer marking process:

Great to be able to compare and see how many different ways the assignment was done. [Student F]
2.e  The collaborative involvement of students in the quality assurance activity.

The performance of the marking key was improved by the collaborative involvement of the students in the quality assurance of the marking key (Activity 2). This involvement of students was only made possible due to the use of ICT that combined online access within the purpose-built building that contained five computers per tutorial classroom. This allowed the students to access the EPSS via the network and record their group comments. Although the equivalent process would be possible in a paper-based form, the cost and management issues would be prohibitive. Although this student involvement occurred only in development cycle 4 (Tutorial Paper), both the students and tutors commented positively on the activity.

Activity 2 entailed the quality assurance of the marking key by the tutors. This was considered by them as vital to the performance gains of the whole task assessment process. As Tutor B commented:

> When we wrote, when we had to say in the actual writing of the rubric, we had ownership… [Tutor B]

When the students were involved in the quality assurance process, an even greater understanding of the assessment task goals and of the desired learning outcomes were gained. A common comment of students when they were not involved in the quality assurance process was that they often did not understand the words used in the grade descriptors. This was noted by Tutor B:

> … I felt that when the students picked up our rubric and read it, because they didn't have ownership, I didn't think that they comprehended it to 100%.

 Assertion 3: The coordinator’s performance was improved and enhanced.

The coordinator’s performance of the task assessment process as a whole was shown by the analysis of the data to have been improved and enhanced in a number of ways. The most important areas of these performance improvements and gains are discussed under the following contributing factors:

3.a  The quality assurance of the marking key.
3.b  The student feedback and reports.
3.c  The management and reporting of the task assessment.
3.d  The control over the confidentiality of student information.
3.e  The output control of the student feedback and reports.
Throughout the study the coordinator was both the coordinator of the unit and the sole lecturer of both units in the Development Phase, which is a typical situation in the School of Education. The roles and responsibilities of a unit coordinator with regard to assessment as listed below (the numbers refer to the numbering system used in the ECU 2003 document on the policy and procedures of assessment (Learning and Development Services, 2005)):

6.2.9 to provide, where appropriate and possible, opportunities for students to participate in identifying their learning needs, planning their learning experiences and determining the ways in which they will be assessed;
6.2.12 to put in moderation mechanisms to ensure consistent marking of all assessment tasks;
6.2.13 to maintain the confidentiality of personal student information including assessment results, except for legitimate University purposes;
6.2.14 to advise the Head of School, Program Director or Course Coordinator as appropriate, of marks and grades by the due date and to attend meetings of the School Board of Examiners;

6.4 Administration of Assessment
6.4.2 Determination of Marks and Grades
6.4.2.d Examiners should ensure that marks and grades are awarded appropriately. Where grades are allocated according to a set of pre-determined standards, students must be provided with the standards prior to completing the assessment.
6.4.3 Reviewing assessment marks and grades for assignments, tests and final examinations.

These roles and responsibilities are referred to in the discussion that follows. The improvements in performance occurred in all seven activities (see Figure 9.1) from the design and development of the marking key (Activities 1 and 2) through the pre- and post-marking moderation, reporting and management processes (Activities 4, 5, 6 and 7). The performance gains in Activities 1 and 2 were due to the processes introduced to improve the quality assurance of the design and development of the marking key, whereas the EPSS contributed more directly to the coordinator’s performance gains in the overall management of the task assessment processes (Activities 4, 5, 6 and 7).

3.a The quality assurance of the marking key.

The quality assurance of the design and development of the marking key enhanced and improved the coordinator’s performance. These processes covered one of the University’s mandated requirements for the coordinator to ‘put in moderation
mechanisms to ensure consistent marking of all assessment tasks’ (6.2.12). Prior to this study, the coordinator had not directly involved the tutors in the quality assurance of the marking key design and development. Nor had she focused as much time and attention on the quality assurance of the marking key as the process developed throughout the study recommends. As a tutor in other units, the coordinator had also identified this lack of tutor involvement in the development of the marking key as problematic.

The team’s responses from the interviews revealed a similar experience to that of the coordinator, that is, previously they were often kept out of the process, with little or no input into the design and development of the marking key. They were left in the dark as to how to interpret the supplied marking key, when supplied one at all. Team members stated that the marking key was often limited in detail and that they were often unsure of what the coordinator wanted assessed in the marking of the assignment. They commented that this lack of detail and vagueness of the marking key contributed to the tutors’ stress and indecision during the marking process. In this case, the tutors’ improved performance in marking due to the early intervention strategies and the EPSS had a flow-on effect and contributed to the coordinator’s improved performance.

Thus, with little or no prior experience of designing and developing an instructional rubric, the coordinator, with the assistance of the researcher, explored and investigated this important area with the aim of improving the performance of the marking key. Both the coordinator and tutors in the interviews identified the importance of the quality assurance process of designing and developing the marking key (Activity 1 and 2) in the overall task assessment process.

The following comments from the coordinator and Tutor E reflect on their past experiences of the task assessment process:

[Usually] a very general marking key, with very little discussion about what the criteria actually meant, limited moderation, ... using very global judgements, like outstanding, satisfactory, not satisfactory, right down to using very specific marks but no discussion. [Coordinator]

Biggest difference for me this year as opposed to years before is the attention to developing good rubrics. ... The collaboration on developing the rubrics is another big difference. [Tutor E]

The coordinator found that a number of important goals in performance were achieved by undertaking the quality assurance of the marking key in collaboration with the tutors. This collaboration often took place three to four weeks before the marking of the assessment task (Activity 5). This early collaboration and involvement of tutors
meant that they had time to devote to the quality assurance process. This also meant that
they had time to process and assimilate the marking key criteria and grade descriptors,
thus reducing the cognitive load before marking occurred. This resulted in the marking
key being more aligned with desired learning outcomes, more explicit and more
transparent (by the use of an instructional rubric design). This was reflected in the
following comments on rubric development and quality assurance by the Coordinator:

The development of the rubric (marking key) helps not only the students’
understanding of what’s being expected of them but also staff, in terms of what
they will be looking for. [Coordinator]

Different approaches were tested to involve tutors in the very early stages of the
design of the marking key. However, time constraints, lack of ICT skills, and technical
issues involving access to the Internet and servers all contributed to the decision by the
team to discontinue these initial attempts. Instead, the coordinator developed a working
copy of the marking key rubric based on an existing rubric selected from the literature,
and the tutors were involved in the quality assurance of the working copy and came to a
consensus on the final version of the marking key. This approach was found to be the
most effective and time saving, both for the tutors and the coordinator, who saved
additional time as she gained experience in this process.

3.b The student feedback and reports.

The coordinator improved quality assurance and control over the student feedback
and reports. This was made possible through the features built into the EPSS. These
features included the ability to carry out spell-checking and readily review the tutors’
feedback comments. Additionally, the analysis of the distribution of student marks and
sub-marks within tutorial groups, between tutorial groups and in the unit as a whole was
also simplified. The built-in ability of the EPSS to sort and find on any field allowed,
for the first time, the analysis of the criteria marks. Thus, a quantitative analysis of the
criteria marks became possible, whereas previously this analysis had been limited to a
qualitative assessment based on tutor feedback – usually provided verbally and often
many weeks after the marking process had been completed.

These features of the EPSS contributed to the coordinator’s ability to better carry
out the University mandated responsibility of having ‘moderation mechanisms to ensure
consistent marking of all assessment tasks’ and of ‘reviewing assessment marks and
grades for assignments, tests and final examinations’ (6.2.12 and 6.4.3 respectively).
These could now be reviewed easily, adjusted and spell-checked uniformly, not in an adversarial way but from a quality assurance viewpoint of monitoring the overall student response to the assessment task.

A feature of the reporting aspect of the EPSS that was greatly appreciated by the team, and specifically the Coordinator, was the access to and easy use of the built-in spell-checker. The spell-checker not only detected spelling errors but also minor grammatical errors, and it gave the reviewer time to edit the comments without scrolling through 200 student feedback pages. This was because the spelling view of the data displayed the comment boxes only on one screen, thus improving the functionality of the spell-checker. This feature, made possible due to the EPSS, was appreciated by both tutors and coordinator for it allowed for the first time a final quality assurance of the feedback comments, and the types of feedback being given. This had the added bonus of giving the coordinator a good sense of how the tutors and students were doing in carrying out the assessment requirements, which was previously difficult to achieve.

Tutor B enthusiastically commented on this feature:

_"Oh, the spell-checker was good. My typing is not up to scratch so the spell-checker was really good. ... Yeah, keeps it all uniform ... It's thorough, it shows evidence that ... it shows that it's thought through, and besides that the students are getting feedback from the rubric itself, the whole thing is very well thought through and planned, and organised, and when you're organised it's easy to sit down and mark. If I feel organised, I can sit down and mark." [Tutor B]_

3.c The management and reporting of the task assessment.

The coordinator’s performance was enhanced by the improved ability to manage, analyse and report on the task assessment activities and results. These activities and results were made easier through the features built into the EPSS, which contributed to the coordinator’s ability to better carry out the University mandated responsibility of having ‘moderation mechanisms to ensure consistent marking of all assessment tasks’ and of ‘reviewing assessment marks and grades for assignments, tests and final examinations’ (6.2.12 and 6.4.3 respectively). The ability of the EPSS to efficiently combine the different tutors’ marks that were imported electronically into the coordinator’s master copy of the EPSS, allowed the coordinator for the first time to have access to the unit results, including sub-marks and comments and not just the final total mark, before they were returned to the students. This allowed the coordinator to analyse the unit assignment marks, sub-marks and comments and carry out an overall
quality assurance where previously this process could only be achieved for the assessment mark.

Previously, there had not been any mechanism available to the coordinator for the efficient collection and collation of assignment marks. This process would have been difficult and time consuming without the application of the EPSS. Therefore, the new ability that the EPSS afforded to the management of activities and the analysis of results improved the performance of the coordinator. These gave the coordinator a much more detailed and accurate view of the students’ results that covered how the tutors were marking, the spread of marks for the tutorial groups and the unit, and the progress of individual students.

For the first time, if students wanted to discuss their progress with the coordinator, it was possible to access the student’s feedback view electronically from the EPSS and thus enter into a meaningful dialogue with the student about their results. Previously, the coordinator could only talk in very general terms before talking with the actual student’s tutor. As most tutors were sessional, it was often difficult for them to meet with their students outside tutorial times. The procedure observed by the researcher and reported by the experts during the Exploration Phase indicated that the coordinator often only saw the students’ total mark for the assignment and in some case only the unit total mark. This left all the responsibility on the tutors to mark fairly and reliably. The range of marks between tutors was often felt to be significant by students due to the lack of transparency in the marking process. The management procedures and activities proposed in this study greatly improved transparency, reliability and management of the marking activity.

The coordinator explained these management performance gains in the following words:

*Very laborious process always up until now for me in terms of collating the marks ... Being able to look at the spread of marks, having that information all in one place was hugely different for me.* [Coordinator]

Thus, the important quality assurance and control of the students’ feedback comments and marks were carried out by the coordinator and, when necessary, in collaboration with the tutors. The issue of how to access and print this information also arose. The EPSS allowed the coordinator not only to retain all this information on the students’ task assessment results in an electronic form, but allowed the data to be displayed in many views. These views of the assessment data could be printed easily
and efficiently when necessary. Typically, during the development cycles, the *student feedback* sheets were sorted by tutorial group and then by student surname, and printed and placed in envelopes for tutors to collect. The coordinator found this process involving approximately 200 students could easily be done within an hour.

3.d The control over the confidentiality of student information.

The application of the EPSS improved and enhanced the control over the confidentiality of student personal information including assessment results, compared with previous methods used. The University specifically mentions the coordinator’s responsibility ‘to maintain the confidentiality of personal student information including assessment results, except for legitimate University purposes’ (6.2.13).

Before using the EPSS, the coordinator relied upon the login to her computer and the locking of her office door as the main security devices to prevent access to student personal information such as assessment results. She also had very little control of how the lecturer or tutors within her unit handled the confidentiality of students’ personal information.

3.e The post-marking moderation and monitoring of students’ results.

This component focuses on the ability of the EPSS to monitor the student and tutorial group across assessment tasks, including the total unit mark. The post-marking moderation processes that the EPSS provided improved the coordinator’s performance and were an enhancement to prior processes. As stated previously, there had been no consistent method for returning tutors’ marks to the coordinator. The coordinator could request they be in spreadsheet format, but many tutors were still novice users of ICT and this could be beyond their skill level. This issue was discussed in the *Exploration Phase* of the study, and it was found that the time constraints at the end of the semester meant that tutors focused on collecting and collating the total student unit mark ready for submission to the Examination Board. Time did not remain for an overall analysis of the individual assessment tasks or students due to the inefficient management procedures in place prior to this study. The focus of the coordinator was on students with borderline marks and involved discussion with the relevant tutors.

The implementation of a criterion- or rubric-based marking key meant that the student assessment marks did not need to follow the University’s recommended marks distribution. This is highlighted by the University guidelines on assessment that
examiners should ensure that marks and grades are awarded appropriately. Where grades are allocated according to a set of pre-determined standards, students must be provided with the standards prior to completing the assessment’ (point 6.4.2.d). This meant that the time consuming and unproductive work of checking that the students’ assessment marks followed the recommended distribution was eliminated. The EPSS was able to show and graph the grade distribution of student assessment marks by assignment, tutorial grouping and final unit total. This allowed the coordinator, for the first time, to quality assure and control the assessment marks and not just the unit total mark as had been done in the past.

The coordinator’s performance was enhanced and improved by her improved ability to analyse the task assessment results. Previously task assessment results had not been readily obtainable for analysis by the coordinator. The implementation of the EPSS changed all this by allowing the coordinator for the first time to monitor not only the results within an assessment task but within each tutorial group and the unit as a whole. The EPSS allowed both the coordinator and tutors to readily find individual students, and review their marks and comments and adjust if necessary. They were also able to easily sort and group by criteria or total marks for the first time within the unit or by tutorial group. No longer was it an estimate about what and how many students were having difficulties with specific criterion within each assignment, now individual students could be identified, contacted and monitored.

This meant that assessment results could now be more efficiently quality-assured and controlled as discussed above; and not just the total mark but the sub-marks and comments could be analysed and reviewed electronically. This information could be used to inform the teaching practice both now and in the future. Furthermore, this information could be used to inform tutors and students about the progress of the students’ learning. Under previous task assessment processes, these features had not been possible, practical or feasible within the unit budget.

**Assertion 4: The tutors’ performance was improved and enhanced.**

The tutors, including the coordinator in the role of lecturer, found that they all had individual task assessment performance gains throughout the Development Phase. This was mainly due to their increased understanding of, and participation and collaboration in, the task assessment process. A secondary performance gain was observed by the
researcher in tutors’ improved ICT skills, confidence in using ICT, and a willingness to
explore and experiment with ICT. In addition, the interest generated by their
enthusiastic involvement in the research study has led to the further features and
recommendations being incorporated into the EPSS in the post-development of the
EPSS and its use in other units on a trial basis.

The tutors’ performance of the task assessment process as a whole was shown by
the analysis of the findings to have been improved and enhanced in a number of ways.
The most important areas of these performance improvements are listed and are
discussed under the following contributing factors:

4.a The timely returning of assignments.
4.b Students’ rights to reasonableness and fairness.
4.c Tutors’ professionalism.
4.d The quality of the feedback to students.
4.e The pre-marking moderation.

As stated in the previous section, a number of roles and responsibilities have been
mandated to lecturers and tutors by Edith Cowan University in relationship to the
assessment process. The following further roles and responsibilities relate directly to the
task assessment process (the number in brackets refer to the numbering system used in
the ECU 2003 document on the policy and procedures of assessment (Learning and
Development Services, 2005)):

6.2.15 to assess students’ work fairly, objectively, consistently and in a timely
manner and to provide adequate feedback about performance;
6.2.16 to provide timely feedback on assessments during the teaching session.
To ensure for the regular semester that turnaround time for
assessments is three weeks or less. Any increase in turnaround time for
assessments requires the approval of the Head of School or Program
Director and students must be advised of the increased turnaround
period in advance;
6.2.17 to prepare and present unit material at an appropriate standard and
within the resources available;
6.2.18 to be available at reasonable times, as approved by the Head of
School, so that all students, whether they are enrolled in on-line mode,
external mode or face to face mode, may discuss aspects of the unit
including learning issues and feedback from assessment excluding
final examination. Staff will advise all students, having regard for their
mode of study, of their availability for student consultation;
6.2.21 to maintain the confidentiality of personal student information
including assessment results, except for legitimate University
purposes;
As mentioned in the preceding section on the coordinator assertion findings, the coordinator in this study took on the role of both coordinator and lecturer. Thus, in analysing the data from the semi-structured interviews, quotes from the Coordinator as they relate to her role as lecturer/tutor will be used.

4.a The timely returning of assignments.

The returning of assignments within the two-week period, as per University requirements, was easier to achieve because the tutors found their ability to complete the marking in the required time was improved and enhanced by the EPSS. This covered two of the prescribed University’s roles and responsibilities, firstly ‘… to ensure for the regular semester that turnaround time for assessments is three weeks or less’ (6.2.16) and secondly to assess students’ work ‘in a timely manner’ (6.2.15). At the time of the study, the prescribed time was two weeks but due to the difficulty of meeting this tight time constraint, the official policy has been amended to three weeks. However, the use of the EPSS did permit the meeting of the previous two-week deadline more easily.

The marking process, Activity 5, was where the results of processes implemented in previous activities and the EPSS came together as an integrated whole. Marking was usually a very busy and stressful time for the tutors. In the interviews tutors agreed, usually strongly, that the EPSS was of great assistance and helped in their performance of the marking process. The EPSS allowed the marking key rubric to be accessed via the computer screen. All interactions with the EPSS, such as the recording of criteria word-grades, movement between screen views, the finding of students’ recording sheets, and revising the grades or comments (or both) could all to be done via the mouse - except for the entry of the tutor feedback and this was via the keyboard.

This method of handling the marking and recording had many advantages but for this assertion, what was decisive was the efficiency and time saving. This was achieved by the elimination of the busy, unproductive and time-consuming administrative and clerical work involved in the marking process. The anxiety, stress and time off-task of this unproductive and clerical work involved in the traditional methods of marking, such as adding up the sub-marks, recording the total and re-entering the marks in list form for submission to the coordinator for each assignment, were eliminated with the use of the EPSS. All the interviewees commented that they especially liked not having to do the busy work, like ‘adding-up’ manually. The automatic adding-up also removed
the possibilities of human error that a number of tutors mentioned as being especially important. The following quotes illustrate this point:

In the past, I was terrible at adding up, I double-checked everyone, I would have been nervous that I had made a mistake. .... The automatic adding up of the marks, the student names being there automatically, just saves so much time. [Tutor E]

That it saved an enormous amount of time in marks allocation and adding up. Which I really value because that’s something I do very slowly. [Tutor F]

Big benefit for me was that it reduced the busy work. [Coordinator]

It (the tool) saved a lot of time because you didn’t have to do all the fiddly figures, and you didn’t have to double-check or triple-check the adding up in case you’d made a mistake because it was all done up automatically. [Tutor B]

Finally, the ability to switch between the different views of the marking data was found to be useful during the marking activity. The grade sub-marks and comments were recorded once and could be viewed in different ways: tutor marking, student feedback, and spreadsheet views. The tutors could instantaneously see the growing spread of marks as they marked individual students. They did not have to re-record the students’ total marks and, if they did decide to change a sub-mark, the total was instantly re-calculated. The use of emotive language to describe this feature (beauty, luxury button) shows Tutor B’s enthusiastic response to the EPSS:

You know, the beauty of it was really that at any stage (of marking) I could go back and have a look at the overall marks and switch between views ... I could be checking that as I'm going and just by clicking buttons, not having to do any adding ... it's quite like a luxury button. ... It's so easy, so much easier. And going back and changing marks, it's so easy to go back and just changing marks. [Tutor B]

Thus, the study found that, irrespective of the type of assessment task and marking key (the interviewees had all experienced at least three types of marking keys), the tutors all agreed that the use of the EPSS saved them a significant amount of time and directly helped them to achieve the goal of returning the marking within the required two weeks.

4.b The students’ rights to reasonableness and fairness.

According to the tutors, the instructional marking key rubric enhanced the students’ rights to reasonableness and fairness in the task assessment process (covered in the policy guidelines 6.2.15, 6.2.16 and 6.2.17). Prior to the use of the instructional
rubric, the students were either not given a marking key or were given one that was just a rewording of the unit outline and contained only the breakdown of the mark allocation and no grade or level descriptors.

The tutors and students found that these types of marking keys made it hard to understand and evaluate what the differences were between the number ranges, for example, what x out of 10 really meant. This contributed to low inter- and intra-tutor marking reliability and thus limited the students’ rights to reasonableness and fairness in the task assessment process. These students’ rights are referred to in policy guideline 6.2.15: ‘to assess students’ work fairly, objectively, consistently and in a timely manner and to provide adequate feedback about performance’.

The implementation of the instructional rubric enhanced the students’ rights to reasonableness and fairness in the task assessment process. A contributing fact was the increase in the transparency that the instructional rubric provided. Team members described these performance gains in the following terms:

*Improve the information you give students before they do the assignment [the marking key]. So, they have a better change of performing well on it. … This instrument [marking key] should also help us be clearer in explaining what we want from students in the beginning and so that miss match of their perception of fairness and what is really fair should decrease as the result of using the instrument because they will be clearer to about what we expect.* [Tutor F]

*The development of the rubric helps not only the students’ understanding of what’s being expected of them but also staff in terms of what they will be looking for.* [Coordinator]

*The actual marking process makes it a lot quicker and easier and a lot fair for the students.* [Tutor E]

4.c Tutors’ professionalism.

Tutors’ professionalism was enhanced through the quality of the student feedback sheet. This assertion covers the policy guideline ‘to prepare and present unit material at an appropriate standard and within the resources available’ (6.2.17). The EPSS enabled tutors to customise the student feedback view, the one printed and returned to the students. Further, it allowed the tutors’ feedback comments to be spell-checked and, if necessary, to be elaborated on later after reflection and before the final printout, which all improved the output quality. The team commented that these features would have been a “very time consuming” process if they had tried to do this in the traditional paper-based marking format. That is, traditionally, students received handwritten marks
and comments that had usually not been reviewed. Thus, the team all felt that the student feedback sheet that EPSS produced and returned to the students, was of a “higher professional standard” than previous or current marking and recording methods and contributed to their feeling of “professionalism”. All the tutors made similar comments to that of Tutor B:

*The other thing about it, is proficiency, I feel proficient when I'm using it, that could be a side issue, but I really think that teaching is a profession and I feel really proficient when I'm using something like that.* [Tutor B]

4.d The quality of the feedback to students.

The tutors’ quality of feedback was improved and enhanced, which covers the policy guideline ‘to assess students’ work fairly, objectively, consistently and in a timely manner and to provide adequate feedback about performance’ (6.2.15). The interviewees thought that the quality of their feedback was improved and enhanced by both the use of the instructional rubric and the use of the EPSS tool that incorporated the rubric. The instructional rubric not only informed and guided both the tutor and student in what was being assessed in the assessment task, but also provided detailed and focused feedback to the student. During the marking process, tutors continued to record comments on the assignment when appropriate and summarised these in the tutor comment box.

A substantial amount of time was saved in the marking process by the use of the EPSS and this contributed to improved and enhanced feedback. Tutors attributed this to being able to focus on the actual marking or judgement and not to be continually distracted by the busy work involved in the marking process. The Coordinator and Tutor F illustrated this in the following quotes, where they highlight their experience of improved quality feedback and professional judgement and time to reflect on their teaching practice:

*I was able to spend the time actually considering the student work, instead of doing the busy work of writing the name on the sheet, writing of the marks, writing the comments. Because the bottom line is you have only x amount of time for each paper that you are marking. So, if your time is taken up doing that busy work you have less time to spend on actually making judgements about the student’s work. So, for me the big difference was that I was able to spend more time doing what I think assessment is about feedback.* [Coordinator]

*It meant that I had extra time to focus on the really important things about marking, about determining quality and thinking about what learning my students were demonstrating. So that them informed my teaching, so what are the gaps here, what do I need to really work on in my classes. What*
Another aspect of the EPSS that also contributed to the quality of the student feedback was the tutor’s global comment box. This feature of the EPSS was the result of the quality assurance process, Activity 2. This feature was added to the EPSS at the request of the team during the collaborative feedback sessions that occurred throughout the development cycles. The global comment box allowed elaborate and detailed generic or common comments to be recorded once by the tutor, and then be accessible on every student tutor marking view. These generic comments could be added to or copied and pasted from the student comment and then edited. They could also be easily copied into the tutor comment box when appropriate. This feature greatly helped with the consistency and quality of the tutors’ feedback to students, as described by Tutor A:

*I think the quality of the comments was a lot better because you could keep the ones that you'd already written and refine them. Sometimes you had, you'd write a very good comment, and in the old way, you'd think 'oh, what was that really good comment I wrote?' Because this is the same as that, whereas by being able to keep it on the computer, you could re-use them and your wording was always spot-on because you'd copy and save your really well-worded one and you'd use them again. So I think the actual quality of the comments, the feedback, was much better. …* [Tutor A]

The Coordinator, reflecting on her use of the global comment box, found that the detailed rubric actually reduced the need for elaborate and detailed feedback.

*Didn’t use it [global comment box] that often, that’s interesting, partly because the rubric actually reduces the number of comments that you needed to put in, the rubric does the comments, so you really only using a fairly generic type of individual comment. [Coordinator]*
adjustment. This final editing, if required, typically related to issues of interpretation and emphasis and did not affect the grade levels, and it was agreed upon collaboratively during the moderation meeting. The EPSS allowed for these last-minute fine adjustments, whereas previously, if a large number of marking sheets had already been printed this possibility would have been lost.

The interviewees felt that the moderation process became more transparent, precise and definitive as the development cycles developed and as the instructional rubric design and development became more precise. The ease of interpretation and marking afforded by the marking key, when combined with the EPSS, contributed to the performance enhancement of the moderation process. The paragraph grade/level descriptors were found to be much more useful during the whole-task assessment process than the more usual number or grade line for each criteria. As Tutor F put it most clearly, the grade descriptors provided a *common language* and neutral space that could be used in the moderation discussion, whereas previously valuable time was wasted in working out for example what 6 out of 10 meant to the different tutors during the moderation meeting:

*It [the rubric] made the [moderation] process easier. It gave us a common language and it gave us the capacity to pinpoint aspects of the criteria, which had been inappropriately weighted. Your writing them up, your sharing them with others, you are putting them out there to be tested by other people.... And I think that's a really healthy process. And it does not necessarily happen unless you have got something neutral for it [moderation] to happen in, like an instrument. ... And making the conversations around that process means, that you're making your judgements, your professional judgements more explicit. [Tutor F]*

Tutor A’s comments further emphasised the significance of the rubric in the moderation process, explaining how the pre-marking moderation process improved her marking:

*[The rubric] just helps the dialogue because it's specific and you've got something to talk about, and it's in that discussion that you find out, oh well everybody expected something more than I expected, so I need to rise my standards. [Tutor A]*


**Conclusion**

In this chapter, the analysis of the data was discussed to address the research question:

> To what extent does the application of an electronic performance support system (EPSS) enhance and support the performance of the task assessment process; the management, reporting, marking key development, marking, feedback and moderation processes where professional judgement is required in the task assessment of student work in a university course of study?

The previous chapter (Chapter 8) focused on the performance improvement and enhancement of the task assessment from the more limited perspective of the developed EPSS. This showed that significant and important performance gains were achievable. However, the importance of other intervention strategies when combined with an EPSS produced a performance synergy significantly greater than any one-intervention strategy applied alone to the task assessment process. These intervention strategies were based on the areas of possible intervention discussed in Chapter 3: organisational systems, cognitive support, and skills and knowledge. Thus, the main intervention strategies developed were an improved management system, an enhanced marking key, and improved knowledge transfer between the coordinator and the tutors involved in the task assessment process.

The focus of this chapter was on the performance improvement and enhancement that combined the synergy of the above strategies in the task assessment process as a whole, and it involved mainly the findings from the summative interviews of the coordinator and tutors involved in the Development Phase. This analysis was grouped under four major assertions. The first two related to the performance gains due to the instructional rubric (marking key), and the last two related to the performance enhancement of the coordinator and tutors. The study found that the intervention strategies led to: the marking key being improved and enhanced; the educative value of the marking key being improved and enhanced; the coordinator’s performance being improved and enhanced; and the tutors’ performance being improved and enhanced.

These assertions, and the contributing components, all addressed different aspects of the research question. All assertions demonstrated that the application of an EPSS combined with the developed intervention strategies could provide significant performance support to the task assessment process involving professional judgement. This performance support was achieved to varying degrees across the seven activities...
identified within the task assessment process. The significant performance support occurred during Activities 5, 6 and 7. However, this support was built upon the support provided in Activities 1, 2 and 4.

The final chapter (Chapter 10) presents: a summary of the study; reflections on research process and the role of the researcher; the major performance improvements achieved through the developed EPSS; the post-development evaluation; the significance of the study and implications to theory, policy and practice; and recommendations for further investigation and practice.
CHAPTER TEN
CONCLUSIONS

This final chapter includes a brief overview of the study and discusses the intervention strategies and activity areas where performance improvement was demonstrated to have occurred in the task assessment process. This chapter also discusses the *Post-Development Phase* of the study resulting from the success of the study that led to the implementation of the developed EPSS and strategies by other unit coordinators, and the recognition both within the local and in the wider academic community. The significance and implications of the study are discussed next, concluding with recommendations for further investigation and practice.

**Study Overview**

This study began with the premise that the performance of the task assessment process involving professional judgement could be improved through the application of an EPSS. The focus of computer-aided assessment (CAA) had been until recently focused on objective-based assessments to the exclusion of subjective-based assessments. This meant that a large number of assessment tasks, specifically those involving professional judgement, were being assessed with little or no ICT involvement.

The task assessment process was defined by the researcher as all activities carried out by staff in assessing student work. This ranged from designing the marking to reporting the assessment results to students and the institution. In defining the activities involved in the task assessment process, the researcher drew on the literature of assessment and performance technology to identify six specific workplace activities where performance intervention might occur to improve performance.

Based on the study’s premise, the research question became the following:

*To what extent could the application of an electronic performance support system (EPSS) enhance and support the performance of the task assessment process: the management, reporting, marking key development, marking, feedback and moderation processes where professional judgement is required in the task assessment of student work in a university course of study?*

A number of factors coalesced to both justify and enable this study. In particular they were: 1) the importance of the task assessment process to teaching and learning; 2)
the fact that the practices of assessment had not kept pace with the changes in educational teaching and learning practices; 3) the identification in the research literature of a significant gap between the actual performance and the desired performance of the task assessment; 4) the increase in the power and availability of ICT to university staff and the under-utilisation of this technology in the task assessment process; 5) the researcher’s interest in both assessment and software development; and 6) the fact that the area of electronic performance support of the task assessment had not previously been investigated. Based on these factors and others, it was deemed that the study was of substantial significance to proceed to the design stage.

The study design presented a number of challenges, two in particular. Firstly, the study environment was not a simulated or trial environment, as is usual in these software development studies, but was situated within an actual work environment that involved the performance of the task assessment process. While this meant the validity of the findings was high, it also meant that the researcher and the team were placed under the rigors of normal work conditions that involved critical professional judgement, time constraints and stress, plus the additional pressure and stress of attending to the study itself. Secondly, the study design involved six iterative development cycles over two semesters to develop a proof of concept of a task assessment EPSS. This design approach was considered essential to the success of the study because only through the active participation of users and an iterative development process within the actual work environment would the findings be valid and significant.

Based on the literature review (Chapter 2) that identified and investigated three different fields of research: assessment, human performance and software development, the study’s conceptual framework (Chapter 3) was developed. The conceptual framework comprised three connected frameworks derived from three separate areas of research literature. The first one, the Human Performance Technology (HPT) framework, provided the overarching theoretical framework in which to investigate, develop and evaluate the performance interventions strategies, and developed out of the literature on performance support (PS). The HPT model was adapted and modified by the researcher into the three phases of the study, the Exploration, Development and Evaluation Phases. The second framework described and positioned the task assessment process within the assessment process and highlighted the task assessment areas where possible interventions to improve performance might occur at the
workplace level. The third and final framework described the software or product design approach to be used to develop the EPSS. This framework highlighted the importance of participatory action research (PAR), and a performance centred design (PCD) approach to the development of a useful and usable product. These frameworks provided the basis upon which the research design (Chapter 4) was refined and developed.

The Exploration Phase (Chapter 5) included a preliminary investigation of the performance gap of the task assessment process through the process of Performance Analysis and Cause Analysis, as described in the HPT model. This was limited by the lack of any quantitative metrics of the task assessment process. This analysis resulted in seven major task assessment activities being identified within the task assessment process where interventions might be applied to achieve performance gains. The seven activities were: 1) marking key design, 2) quality assurance of marking key, 3) incorporation of marking key within the EPSS, 4) pre-marking moderation, 5) marking, 6) post-marking moderation, and 7) management.

The constraints and limitations of the study only allowed a partial combining and blending of the interventions within, and between, these activities and the integration of these seamlessly into the evolving EPSS, over the six development cycles over two semesters (Chapters 6 and 7). An analysis (Chapter 8 and 9) identified two pivotal activities in the complex task assessment process: the marking key design and quality assurance of the marking key, both of which needed to be accomplished to a high standard for the assessment process as a whole to achieve its many goals. Further to this, the study demonstrated that quality assurance processes need to be implemented within every activity to guarantee a high performance standard.

**Reflections on research process and role of the researcher**

The study involved three distinct phases: the Exploration, Development and Evaluation Phases that required different methods of investigation. Due to these different methods, the type of data collected and analysis was also different. This meant that the researcher’s role changed throughout and even within each phase of the study. In the Exploration Phase, the role of the researcher involved the collection and analysis of data on the desired and actual performance of the task assessment process within the environment that the study was to be undertaken. This involved semi-structured...
interviews of assessment experts, observations of the task assessment process and
document analysis. The data was used to identify the performance gap and potential or
possible intervention strategies that might be employed within the scope of the study to
reduce the gap. On reflection this phase of the study although only explorative in nature
might have been improved by including a survey of staff within the School of Education
on the use of ICT and types of marking keys used in their task assessment process and
where possible to obtain copies of their marking keys. This would have provided a more
comprehensive view of what was happening within the School with regard to the task
assessment process.

The second phase, the Development Phase of the study involved the UCD method
of software development in a live working environment through six iterative
development cycles over one year with the role of the researcher being markedly
different from the previous phase of the study. The role of the researcher in this phase
of the study was that of an observer, recorder and programmer for the team members
design instructions during the task assessment process. These were subordinate roles
within the team and involved implementing recommendations regarding the design and
development of the EPSS and the intervention strategies to improve the performance of
the task assessment process. The live working environment placed both the developer
and team members under additional stress and meant that unlike in a simulation or test
environment failure or breakdown was unexcitable. In addition, this meant that the
researcher or team members could not take time-out to reflect and discuss issues.

On reflection on this phase of the study, the researcher found that because of the
unique nature and scope of this phase it was hard to comment and make suggestions on
how it could be done differently. Although the process could be replicated, the skill set
of the researcher and team members would be difficult to match. The use of the live
working environment needs a great deal of commitment by both the researcher and the
team members. The timeline line between getting the design brief from the team and the
implementation by researcher was often very short. This meant that often the researcher
was carrying out a number of roles simultaneously. Reflecting on this the researcher
considered that although having one person doing all the research roles was ideal for
development of the EPSS, the reality was that the commitment was substantial and
would be difficult to justify again.

During this phase, the use of the Internet as a collaborative tool was possibly a
little too early with the limited resources available to the researcher and meant that it
was abandoned, however, now with the availability of web2 applications the collaborative aspect of the Internet would be worth investigating. Clearly, with a change in the scope of the study, for example, more resources then changes could be imagined. For example, task observations of team members using the EPSS could have been videoed to allow for detailed analysis. In addition, there have been many changes in the technology field that would impinge on both the design and implementation of the study.

In the final phase, the Evaluation Phase, the role of researcher was, as for the first phase, again one of data collection and analysis. Data collection involved the semi-structured interviewing of the individual team members involved in the Semester 2 of the Development Phase. The researcher on reflection considered that, if time permitted, the interviewing of staff involved in the Post-Evaluation Phase would have been beneficial to the study, as these users had used the EPSS without any training.

Finally, reflecting on the success of the study as a whole the researcher felt that this was due to the combination of the researcher’s professional knowledge and background and the team’s knowledge, willingness and enthusiasm to participate in this demanding study. This combination of participants would be difficult to replicate. The professional experience and qualifications of the researcher in the field of education and software design and development allowed for not only the conceptualisation of the study but the development and implementation of the EPSS. The development and evaluation of the EPSS through an iterative prototype development process using a UCD approach, over six development cycles within a live working environment, while rare, added to the validity and reliability of the findings. Although the team members were mainly novice users of ICT and voluntary participants, they were enthusiastic collaborators and continued suggesting features and strategies to improve the task assessment process. To replicate this situation would probably require a small team of dedicated developers and researchers to work with a team of academics who were provided with some form of extrinsic motivation to be involved.

**Areas of performance improvement**

Organisational systems, cognitive support, and skill and knowledge were identified during the development of the conceptual framework (Chapter 3) as possible areas of intervention to improve performance that were within the scope of the study. The task assessment conceptual framework identified from the literature four key
performance activities: marking, the marking key design, the skills and knowledge of the assessor, and the management and administration within the task assessment process. These key performance activities were further refined into seven activities of the task assessment process. These activities were identified from the research literature, expert interview data and university documents. From this investigation carried out during the Exploration Phase of the study (Chapter 5), significant areas of low performance were identified and resulted in suggested intervention strategies.

Due to the limited development time of one year, the initial focus, as the HPT method recommends, was on the improvement or refinement of existing processes or strategies not the development of new ones. The EPSS was integrated and blended into the activities to varying degrees during the development cycles, with the focus on the marking activities (Activities 4, 5, 6 and 7). These activities involved moderation, marking, feedback, reporting and management, and were the activities that showed the largest gaps in performance during the Exploration Phase. As the study progressed through the six development cycles of the Development Phase (Chapter 6 and 7), the iterative and formative findings began to highlight the importance of the instructional marking key rubric that was designed and developed in Activities 1 and 2. These activities proved crucial to the performance gains of the marking activities that followed (Activities 4, 5 and 6). This unique combining and blending of the instructional rubric design of the marking key, the quality assurance of the developed marking key, and the incorporation of the marking key into the EPSS produced a synergy of performance improvement beyond that of any single one of these strategies.

The performance gains due specifically to the EPSS were discussed in Chapter 8, and were related mainly to improved usability. This improved usability was achieved mainly through the elimination or removal of areas of poor performance. For example, the application of the EPSS removed or reduced much of the busy and unproductive clerical work involved in the task assessment process such as entering student details and marks, adding up and multiple-handling the data. In the area of reporting and management, improvement in performance was also demonstrated.

Performance gains that resulted from the combined usability of the EPSS and the design and use of the marking key rubric were discussed and described through the four assertions discussed in Chapter 9. The four assertions were that:

1) The marking key was improved and enhanced;
2) The educative value of the marking key was improved and enhanced;
3) The coordinator’s performance was improved and enhanced; and
4) The tutors’ performance was improved and enhanced.

These assertions identified improved performance through improved fairness, consistency, validity, reliability, transparency, feedback, reporting and management.

**Post-Development Phase**

At the completion of the one-year collaborative *Development Phase* of the study in 2003, which involved more than 400 first-year students, and a team of six tutors and the coordinator, the researcher was gratified when the team and especially the coordinator wanted to continue to use the developed *proof-of-concept* EPSS and strategies. The developed EPSS, although not yet developed to a commercial level, was sufficiently developed for use with limited assistance. When other lecturers heard about the project through presentations, recommendations and *word of mouth*, they realised the potential of this innovative use of ICT to task assessment and wanted to trial this newly developed EPSS in the task assessment process. This *ripple effect* resulted in many coordinators taking up different parts or strategies of the developed EPSS. This self-selected small group, some of who were already exploring the possibilities of ICT in their work environment, became part of the *Post-Development Phase* that formed part of the confirmative evaluation of the study and is discussed below.

By the end of 2005, nine units, including the original two from the study, had incorporated the EPSS into the task assessment process, with nearly 2,000 students and more than 40 tutors being involved (see Table 10.1). This endorsement of the study’s findings that the performance of the task assessment process could be improved through the application of both the strategies and the EPSS developed was achieved without any promotion by the research, and required no formal training of the staff involved. These results highlight the significance of the application of the HPT model and the PCD approach that were selected for the design study and the decision to develop the EPSS within a real work environment.

<table>
<thead>
<tr>
<th>Year</th>
<th>Semester</th>
<th>Number of Students (≈)</th>
<th>Number of Tutors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>1</td>
<td>200</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>200</td>
<td>5</td>
</tr>
<tr>
<td>2004</td>
<td>1</td>
<td>356</td>
<td>7</td>
</tr>
</tbody>
</table>
The value of this innovative application of an EPSS to the task assessment process has also been recognised both within the university and in the wider academic community, through awards, funding, invitations to present to colleagues at conferences, workshops and through publications (see Appendix L). Some of the awards received by the researcher include the 2006 Citation for Outstanding Contributions to Student Learning from the Carrick Institute for Learning and Teaching in Higher Education; and the 2005 Edith Cowan University Vice Chancellor’s for Excellence in Teaching Citation Award.

**Significance and implications**

The implications of the study for practice, theory and research in the area of task assessment are significant and open up new fields of investigation. This study has demonstrated that support and improvement in performance of the task assessment can be achieved through the application of an EPSS. The application of an EPSS can also improve and enhance the utilisation of the substantial ICT infrastructure available to staff and students, which had previously been used mainly for teaching and learning, but not within the assessment area. This type of EPSS applied to the assessment process has the potential to record seamlessly, and with little additional cognitive load, quantitative performance metrics previously unobtainable, such as time spent marking each student (this could be compared between the whole unit and between tutors), or the number of times a student mark sheet is revisited. Furthermore, the comparison between student grades and time spent marking can now be collated and evaluated electronically.

The developed EPSS now provides many educative assessment opportunities for both teaching and learning, and the potential to evaluate theories about peer and group assessment that were previously too difficult and/or economically beyond the resources of the coordinator. The EPSS provides opportunities for the assessment process to be more than assessment of learning, and become assessment for and as learning. This was achieved in part by the automation of a number of task assessment activities through the elimination of the complex and time-consuming clerical and administrative activities usually associated with the task assessment process. The incorporation of the instructional marking key rubric also provides opportunities for assessment to become
more educative. These performance gains in time and efficiency gave staff the opportunity to spend more time on the educative aspects of assessment, such as feedback, feed-forward and evaluation of teaching and learning.

The developed EPSS has been shown to facilitate the recording and collation of self- and peer-marks and to provide student and tutors access to them without any additional cost in time or effort on the part of the coordinator and tutors. This ability to easily collect and manage the marking activity has enhanced the performance of the moderation, marking and management processes involved in feedback and feedforward strategies.

The use of such an EPSS also presents many opportunities to involve students in the task assessment process, and for the first time the possibility to link the students’ feedback both between assessment items of one unit (demonstrated in this study) and to other future units. This means that future coordinators would not only have access to previous assessment marks but to the complete feedback sheet and as the document is in electronic form, this would take only seconds to access. Previously only the unit total mark would be available to staff.

The EPSS could now be easily enhanced to provide students with their feedback sheet in electronic form. Further, a hard copy or email at the end the semester, or the start of the next semester, could provide a summary of the previous assessment comments as additional feedback to the student. A further enhancement could be to include an overall tutor feedback comment box for the semester. This comment box would summarise the previous two feedback comments and suggest what further work needs to be done for next semester, and ask tutors to record how they have carried out these suggestions. The students’ comments would then be recorded electronically for easy access by the next group of tutors.

The use of the EPSS now gives easy access to a wide range of marks that previously had been unobtainable or difficult to access and that can be investigated electronically without any further data entry. This ranges from marks within an assignment that is sub-marks, and marks between assignments both within the same unit and between units. Time did not permit the study to explore in any depth this important area of analysis, feedback and informed reflection on assessment and teaching across units. However, the possibilities and potential of this new area of investigation have been opened up with the use of this EPSS in the task assessment area where professional judgement is involved.
**Recommendations**

The digitisation of the task assessment process of which this developed EPSS is only an example, is part of the overall evolving and growing digitisation of the educational environment. The possibility of combining or linking the digital form of the task assessment with that of the student task in digital form has been brought much closer as a result of this study. The developed EPSS has demonstrated the feasibility, desirability, and possibilities offered by the digitisation of the task assessment process where professional judgement is involved. This demonstration now needs to be taken to the next stage of development, that of commercialisation. Factors such as cost, user and institutional acceptance, and application to different types of methods of assessment will now need to be investigated.

Finally, technologies such as wireless networking, tablet PCs, and improved Internet access and speed, open up further areas of possible investigation of this type of EPSS and its application to the task assessment process. These technologies could support more student involvement in the task assessment processes, and more tutor and expert involvement in the quality assurance and quality control of the activities involved in the task assessment process. Students could, for example, be given the opportunity to assess their progress on an ongoing basis throughout the task activity by self-marking on a regular basis (e.g., weekly). This could promote in the learner increased independence and self-management skills.

**Conclusion**

This study showed for the first time that an EPSS can improve performance when applied to the task assessment process involving professional judgement. Performance gains were achieved in productivity, usability and improved educational outcomes such as validity, reliability, transparency and consistency, regardless of the method of assessment involved. Although the developed EPSS has been applied across many units and years of studies, it has been limited to teacher education units within one institution. Further research studies needs to be done to validate the findings of this innovative research across different fields of study and institutions. This could include expanding the application of the EPSS to include both primary and secondary education establishments.
This study showed that research into assessment has tended to be focused the assessment task rather than the task assessment process. However, the study demonstrated that research into this area could be rewarding for assessment practice. Quality control and assurance of the whole task assessment process was demonstrated to be critical to performance and to its educative value of the process.

The application of the developed strategies and the EPSS to the task assessment process was shown to improve significantly the performance of the task assessment process. The implementation of these strategies and the EPSS allowed for the first time the testing of theories and practices of assessment that previously had been too time-consuming or expensive to carry out in a rigorous and reproducible way. With the application of an EPSS, the following becomes now possible: to further investigate self- and peer-assessment; to refine the management of students’ feedback, including the ability for this feedback to be emailed to students or accessed via the Internet; to improve the monitoring of students results both at the micro (within the assessment task) and macro level (between assessment task and units); and to obtain improved qualitative and quantitative metrics of the task assessment process.

Finally, the study demonstrated that qualitative or subjective-based task assessment does not need to be sacrificed to objective-based CAA in the name of efficiency and cost saving. An EPSS that combines knowledge management in the form of an educative rubric can improve and enhance the performance of the task assessment process by building on the strengths of the assessor and that of the EPSS. The EPSS has liberated the assessor from the busy and unproductive clerical work associated with the task assessment process and, thus, the assessor can now focus their time on the critical aspect of judging student work and providing meaningful and educative feedback.
REFERENCES


References


strategies for the on-line class: from theory to practice (pp. 110). San Francisco: Jossey-Bass.


APPENDICES

Appendix A – Summary of Data Sources .............................................................. 253
Appendix B – Previous marking key (DC1)........................................................... 255
Appendix C – Oral presentation question (DC1).................................................... 257
Appendix D – Poster guidelines (DC2)................................................................. 259
Appendix E – Poster marking key development document (DC2).......................... 261
Appendix F – Essay assessment criteria (DC3)...................................................... 263
Appendix G – Previous marking key (DC3) .......................................................... 265
Appendix H – Working copy of the essay marking key (DC3).............................. 267
Appendix I – Task requirements for tutorial paper (DC4).................................... 269
Appendix J – Event analysis map instructions (DC5).......................................... 271
Appendix K – Semi-structured interview questions ............................................ 273
Appendix L – Papers and presentations .............................................................. 275
**Appendix A – Summary of Data Sources**

<table>
<thead>
<tr>
<th>Source of Data</th>
<th>Exploration Phase</th>
<th>Development Phase</th>
<th>Evaluation Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DC1</td>
<td>DC2</td>
<td>DC3</td>
</tr>
<tr>
<td>Semi-structured interviews</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experts</td>
<td>6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Coordinator</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tutors</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Meetings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formal</td>
<td>-</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Informal</td>
<td>-</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Student Survey</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Structured Task Observation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tutors</td>
<td>5</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>Literature Survey</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Document Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECU</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Other universities</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: The Development Cycle (DC) meetings often overlapped and occurred during the same team meeting.
Appendix B – Previous marking key (DC1)

EDL 1101: Learning and Development 1

Feedback on Assignment 1: Oral Presentations

Name: _______________  Week: ____ Group: __________

Content
• Logical development of argument and persuasiveness of that argument.
• Shows an understanding of the concepts involved.
• Shows an ability to apply these concepts to the topic under discussion.

Comments:

Mark:      /20

Method
• Ability to capture and hold audience interest.
• Use of visual aids eg. Overhead (not overloaded with information, legible, and relevant), class handouts, audio-visual material such as videos, tapes, posters etc.
• Keeping to time limits.

Comments:

Mark:      /10

Lecturer: ___________  Total Mark:      /30
Appendices 257

Appendix C – Oral presentation question (DC1)

- Part A: Describe something you can do or have learned now that you couldn't do a year ago, or did not know about. Describe how you came to be able to do/know this new skill or concept. What were some of the steps or factors involved in this learning?

- Part B: Ask one child, either a 5-year-old or an 11 or 12-year old, and an older person between 40 and 95 from their community the above questions. Within the team, ensure that the age ranges are all addressed.

- Part C: Analyse the data and discuss the similarities and differences about the way people perceive their learning.

The students were also given the following guidelines for planning an oral presentation through a step-by-step process that is sometimes called the PLAN process (Adapted from (Gibson & Hodgetts, 1990)):

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>ISSUES TO CONSIDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Purpose</td>
<td>What is the purpose of the presentation?</td>
</tr>
<tr>
<td></td>
<td>What results do I want to achieve?</td>
</tr>
<tr>
<td></td>
<td>How should I structure the presentation so I stay on track?</td>
</tr>
<tr>
<td>• Logistics</td>
<td>What physical facilities are available?</td>
</tr>
<tr>
<td>• Audience</td>
<td>What is the audience’s level of knowledge?</td>
</tr>
<tr>
<td></td>
<td>What do they need?</td>
</tr>
<tr>
<td></td>
<td>How can I keep them interested?</td>
</tr>
<tr>
<td></td>
<td>How can I make it easy for them to follow?</td>
</tr>
<tr>
<td>• Nonverbal communication</td>
<td>What visual aids can I use?</td>
</tr>
<tr>
<td></td>
<td>How can I reinforce my talk through positive body language?</td>
</tr>
</tbody>
</table>
Appendix D – Poster guidelines (DC2)

a) Your poster should be able to be displayed in an area of one metre x one metre, and any text must be able to be read from a distance of one metre.

b) In the centre of the sheet, write a list of bullet points that capture your current understandings about how people learn.

c) Select 4 of the theorists that you have drawn from to construct your understanding about learning.

d) On the rest of the poster use visual representations (photos, drawings, collages, icons, sculpture, for example) to show the links between your statements and the theories on which you have drawn.

The students were also given following general guideline about poster presentation in the unit outline:

- What are poster presentations? Visual, public, method of presenting information/ideas/understandings
- Why do poster presentations? Useful teaching strategy Develops skills for professional presentations
- How do you develop a poster? What do you want the audience: to understand? to remember? Critical ideas Fundamental concepts Relationships between them
- How do you present the physical poster? Gain the audience’s attention Good use of diagrams Using colour Key words Pictures/sketches
Appendix E – Poster marking key development document (DC2)

EDL1101-poster presentation
This assignment is designed to demonstrate your present understanding about how people learn, and the theorists who have contributed to your understandings.

<table>
<thead>
<tr>
<th>Criteria 1</th>
<th>Pass</th>
<th>Cr</th>
<th>D</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poster Layout</td>
<td>Pass/fail</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poster less than or equal to 1 metre x 1 metre, and any text must be able to be read from a distance of 1 metre.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria 2</th>
<th>Pass</th>
<th>Cr</th>
<th>D</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>How people learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In the centre of the sheet, write a list of bullet points that capture your current understandings about how people learn.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria 3</th>
<th>Pass</th>
<th>Cr</th>
<th>D</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Theories</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Select 4 of the theorists that you have drawn from to construct your understanding about learning.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria 4</th>
<th>Pass</th>
<th>Cr</th>
<th>D</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of visual representations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use visual representations (photos, drawings, collages, icons, sculpture, for example) to show the links between your statements and the theories on which you have drawn.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria 5</th>
<th>Pass</th>
<th>Cr</th>
<th>D</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catches viewer's interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria 6</th>
<th>Pass</th>
<th>Cr</th>
<th>D</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conveys the central ideas and provides key details in ways that show an understanding of the concepts involved.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criteria 7</th>
<th>Pass</th>
<th>Cr</th>
<th>D</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriate level of detail.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Other possible criteria headings**

- **Layout**
- Presentation eg Graphics, colour, etc
- Text size and colour
- Writing
- Quality of information
- Titles and subtitles
Appendix F – Essay assessment criteria (DC3)

- Does the introduction clearly outline:
  - What the essay is about
  - How the topic will be developed
  - Why readers should be interested?

- Discussion of the topic:
  - The extent to which important issues are raised
  - The extent to which connections are made to theories and literature.
  - In light of the evidence presented in the essay, how reasonable is the conclusion?
  - The extent to which conventions for correct referencing have been consistently followed.
  - The extent to which appropriate grammar, spelling and punctuation conventions are used, and
  - Correct word length.
Appendix G – Previous marking key (DC3)

### EDL1101 ASSIGNMENT 2
### MARKING GUIDE

**Student Name:**  
**Total mark:** /30

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Extremely Well demonstrated</th>
<th>Well demonstrated</th>
<th>Demonstrated</th>
<th>Not Demonstrated</th>
</tr>
</thead>
</table>
| **Part 1** Mark /12 3 resources relate to chosen topic/thesis  
Quality of annotations:  
- Appropriate detail  
- Comprehensiveness: includes main ideas  
- Premise/argument/intent of resource is clear  
- No personal appraisal  
Research process: articulation of processes involved in locating resources | | | | |
| **Part 2** Mark /12 Analysis:  
- Interpretation of 3 authors' arguments/perspectives/positions on topic  
- Comparison of how topic/concept is treated by different authors over different eras | | | | |

**Assignment presentation Mark /6**  
- Referencing conventions  
- Language e.g. sentence structure, punctuation, spelling  
- Word count: 2000 words

**Comments**
# Appendix H – Working copy of the essay marking key (DC3)

<table>
<thead>
<tr>
<th>Tutor</th>
<th>Date</th>
<th>Student</th>
<th>Mark out of 40%</th>
<th>High Distinction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guide to marks</td>
<td>Pass</td>
<td>Credit</td>
<td>Distinction</td>
<td>Clearly and comprehensively outlines the purpose and topic for discussion and how this will be developed.</td>
</tr>
<tr>
<td>Introduction</td>
<td>Outlines some aspects of the purpose and topic for discussion.</td>
<td>Outlines most aspects of the purpose and topic for discussion.</td>
<td>Outlines all aspects of the purpose and topic for discussion.</td>
<td>2</td>
</tr>
<tr>
<td>Comment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discussion; How people learn</td>
<td>Some significant understandings about learning:</td>
<td>An adequate range of significant understandings about learning:</td>
<td>A comprehensive and significant range of understandings about learning:</td>
<td>9 - 10</td>
</tr>
<tr>
<td></td>
<td>Most arguments unsupported with only own opinions stated.</td>
<td>Some arguments substantiated with reference to:</td>
<td>Arguments or explanations *examined and explained from a number of viewpoints.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Where arguments are supported, relies on literal re-statement of course material content.</td>
<td>*other points of view.</td>
<td>*substantiated from a range of relevant literature and/or experience.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*course material</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*some additional appropriate literature</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>*experience.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conclusion</td>
<td>Conclusion provides an inadequate summary.</td>
<td>Conclusion provides an adequate summary.</td>
<td>Conclusion summarise main arguments effectively.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conclusion summarises main arguments effectively and shows critical awareness of the learning process.</td>
<td></td>
</tr>
<tr>
<td>Written communication</td>
<td>Satisfactory standard of presentation and complies with basic academic writing conventions.</td>
<td>Good standard of presentation and complies with most academic writing conventions.</td>
<td>Very good standard of presentation and consistent compliance to academic writing conventions.</td>
<td></td>
</tr>
<tr>
<td>*paragraphing</td>
<td></td>
<td></td>
<td>Excellent standard of presentation and compliance to sophisticated academic writing conventions.</td>
<td></td>
</tr>
<tr>
<td>*topic sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*logical sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*sentence structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar</td>
<td>Presentation has three or more grammatical errors.</td>
<td>Presentation has two or less grammatical errors.</td>
<td>Presentation has no grammatical errors.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spelling</td>
<td>Presentation has three or more spelling errors.</td>
<td>Presentation has two or less spelling errors.</td>
<td>Presentation has no spelling errors.</td>
<td></td>
</tr>
<tr>
<td>Word length</td>
<td>Incorrect word length</td>
<td>Correct word length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intext Referencing</td>
<td>Many errors</td>
<td>Few errors</td>
<td>Minor or no errors</td>
<td>2</td>
</tr>
<tr>
<td>Reference List</td>
<td>Many errors</td>
<td>Few errors</td>
<td>Minor or no errors</td>
<td>2</td>
</tr>
</tbody>
</table>
Appendix I – Task requirements for tutorial paper (DC4)

Task
A written paper (1000 words).
The paper will be peer assessed in the week 5 Tutorial.
The peer assessments will then be moderated by your tutor.

Due date: By 10 am Tuesday 19th August.

Topic:
1. Think of an example from your own schooling experience or your teaching practicum where a child’s learning was directly influenced by the social/emotional dimensions of the situation. The incident/situation you choose may be one in which a child learned effectively, or where his/her learning was limited by the social, emotional aspects of the situation. The situation you describe may be one where the child may have succeeded in spite of difficulties, or where the social/emotional dimensions supported their learning.

2. Write the story of what happened from the child’s perspective. Call this child A.

3. Select one other person who was involved in the event. Call this other participant, person B. Briefly rewrite the incident in the first person as if you were person B.

4. Provide a summary of the conclusions you have drawn from this incident, regarding the importance of the social/emotional dimensions of learning.
Appendix J – Event analysis map instructions (DC5)

1. Present your event analysis map to your small group. Some possible examples may include:
   - Students creating dramatic presentations
   - Teacher led discussion
   - Children engaged in play
   - Teacher instructions followed by assigned work
   - Worksheet task
   - Individual research around a topic of the students own choice and design
   - Presentation to peers.

2. Use a concept map technique to present your information on an A3 sheet of card or paper. Your map should show the relationship between the situation you present, and the learning that is occurring in the situation. Your ‘event analysis’ map may include aspects such as:
   - Social aspects of the learning event
   - Student involvement in learning
   - Locus of control, decision making processes
   - Relationship to theories of learning
   - What learning you think is occurring in this situation

3. The ‘event analysis’ map will be presented to a small group in the tutorial, and handed into the tutor at the conclusion of the presentation for marking.

4. Assessment Criteria
   - Learning event is presented effectively using appropriate multi-media
   - Concept map demonstrates understanding of the links between ideas
   - Learning is clearly identified
   - Theories are linked appropriately

Think of an example from your own schooling experience or your teaching practicum where a child’s learning was directly influenced by the social/ emotional dimensions of the situation. The incident/ situation you choose may be one in which a child learned effectively, or where his/ her learning was limited by the social, emotional aspects of the situation. The situation you describe may be one where the child may have succeeded in spite of difficulties, or where the social/emotional dimensions supported their learning.

Write the story of what happened from the child’s perspective. Call this child A. Select one other person who was involved in the event. Call this other participant, Person B. Briefly, rewrite the incident in the first person as if you were person B.

Provide a summary of the conclusions you have drawn from this incident, regarding the importance of the social/ emotional dimensions of learning.
Appendix K – Semi-structured interview questions

e-Assessment

End of Semester 2 2003 Interview

Name: ______________________________      Date:     __/___/ 03

Thank you for taking part in this research project.

This interview will cover the following areas:

The Marking Process:

• How have you marked and recorded marks previously?
• Was the e-Assessment process different? – Explain
• What were the benefits of using the e-Assessment process?
• What were difficulties and problems using e-Assessment?
• Could you suggest any modifications and improvements to the EPSS and/or process?
• What constraints or obstacles did you find in using the EPSS?
• If you had a choice between using the EPSS or the current method of marking, which process would you choose and why?

Moderation Process:

(Note: The EPSS was used to a limited extent in the moderation process.)

• Would you use the EPSS in its moderation mode where only after marking a piece of work you would be able to see the coordinator’s marks and comments?
• Could you suggest ways in which the EPSS could be used in the moderation process?

Development of Marking Key (rubric):

(Note: The EPSS was used to a limited extent in the development of the rubric.)

• Could you see ways in which the EPSS could aid the collaborative process in developing the marking key?

Suggest other uses or benefits of using the EPSS and process:

E.g.: recording of feedback comments for the coordinator about the assignment; used to review the assessment for next time the unit is run; analyse sub-marks within the assignment; group marking of same assignment on the network (instant display of other markers marks)
Appendix L – Papers and presentations


