Science and Mathematics Education Centre

Online Assessment System with Integrated Study (OASIS) to Enhance the Learning of Electrical Engineering Students: An Action Research Study

Christopher Raymond Smaill

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University of Technology

September 2006

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

Abstract

World-wide, there has been a large increase in tertiary student numbers, not entirely matched by funding increases. Consequently, instructors are faced with large, diverse classes, and find themselves struggling to provide adequate assessment and prompt feedback, two quantities critical in an effective learning environment. Personal computers and the Internet can help solve this problem.

The aim of this study was to develop, implement and validate a Web-based software package that, through providing practice and assessment opportunities, improved student learning and reduced marking and related mundane aspects of instructor workload. At the start of the study, such a package already existed in prototype form: OASIS (Online Assessment System with Integrated Study). As the study progressed, this software package was first fully rewritten and then repeatedly modified.

OASIS delivers individualised tasks, marks student responses, supplies prompt feedback, and logs student activity. Staff can deliver sets of practice questions and assessments to students: assessments may involve different questions for different students, not just numerically different versions of the same questions.

Given my role as teacher, the traditional research ideal of observing without affecting the research environment was both impossible and unconscionable. In particular, since preliminary evidence suggested that OASIS did enhance student learning, I could not adopt a 'two groups' approach to the research, with one group using OASIS while the other did not. Instead, an action research methodology was seen as most appropriate for my double role of teacher and researcher. This methodology enabled me, in the light of my findings, to continuously modify the learning environment and enhance student learning. The action research proceeded through a spiral of one-semester cycles of planning, acting, observing and reflecting.

To maximize rigour, the research ran through eight cycles over four years and involved considerable triangulation. OASIS itself collected much quantitative data. Further data were collected via interview, survey, email and informal discussion from three groups: current students, postgraduates and academics. My colleagues

provided alternative perceptions and interpretations, as did Physics Department academics who were using OASIS, and an external academic who interviewed academics and investigated the implementation of OASIS.

Perhaps surprisingly, academics had generally adopted OASIS to promote student learning rather than to decrease their own workloads. In some cases workloads were reduced; however, where OASIS assessments augmented rather than replaced existing traditional assessments, workloads actually went up slightly. All instructors who used OASIS reported enhanced student learning and wished to continue using it.

Student surveys, interviews, focus-group discussions and informal feedback showed that students found the software easy to use and considered that it helped them improve their skills and understanding. OASIS questions were preferred over textbook questions. Students commonly requested OASIS to be available in more of their areas of study. In general students wanted hints or model answers though some argued against their provision.

The majority of students were enthusiastic about the use of OASIS for practice, and activity logs revealed that they did use OASIS extensively. These logs also revealed the motivating power of assessments: typically half the online practice activity took place in the last 36 hours prior to assessments. Interviews provided further interesting insights into the ways different students approached their studies and assessments.

However, students did voice concerns about the validity of OASIS assignments, noting their peers could rely on the efforts of others to score highly in these. A number of steps were carried out in an attempt to defuse these concerns, including: disabling OASIS practice during assignments, basing assignments on previously unseen questions, and providing different assignment questions to different students.

While this study has achieved the goal of developing, implementing and validating OASIS, many future opportunities exist. OASIS may be used in schools as well as universities. Non-numerical questions, where answers may be somewhere between right and wrong, are possible. OASIS can also be used to deliver concept inventories to students to support research into concept acquisition and retention.

Acknowledgements

I gratefully acknowledge the undergraduates, postgraduates and fellow staff of the Department of Electrical and Computer Engineering at the University of Auckland who were so willing to make time to share their thoughts and experiences with me. I appreciate their input to my study and to the development and implementation of OASIS. In particular, I acknowledge the huge contribution made by Colin Coghill, the OASIS programmer, and also the sterling work of the other members of the OASIS Committee: Mark Andrews, Abbas Bigdeli, John Boys and Gerard Rowe. Their wisdom and experience has been much appreciated and I have enjoyed working with them.

I thank my wife Gail, my parents Joan and Graham, and my children Francis and Simon for their whole-hearted understanding and support throughout my odyssey of part-time study.

I also thank Professor David Treagust, my supervisor, who guided me well on my journey. In spite of many demands on his time, his advice and analysis was always expert and unstinted.

Table of Contents

ABSTRACT ..... **III**

ACKNOWLEDGEMENTS ..... **V**

TABLE OF CONTENTS ..... **VI**

CHAPTER 1. INTRODUCTION **1**

1.1 PURPOSE AND OUTLINE OF CHAPTER 1
1.2 BACKGROUND 1
1.3 CONTEXT OF STUDY 6
1.4 STUDY RATIONALE 10
1.5 RESEARCH GOALS AND QUESTIONS 12
1.5.1 First research goal 13
1.5.2 Second research goal 13
1.5.3 Third research goal 13
1.5.4 Fourth research goal 14
1.6 BRIEF DESCRIPTION OF THE RESEARCH DESIGN 14
1.7 SIGNIFICANCE 16
1.8 LIMITATIONS 18
1.9 SUMMARY 19

CHAPTER 2. REVIEW OF THE RELATED LITERATURE: ASSESSMENT **20**

2.1 PURPOSE AND OUTLINE OF CHAPTER 20
2.2 INTRODUCTION 20
2.3 SUMMATIVE ASSESSMENT 21
2.4 FORMATIVE ASSESSMENT 23
2.5 THE AFFECTS OF ASSESSMENT ON STUDENTS 29
2.6 FREQUENCY OF ASSESSMENT 35
2.7 ASSESSMENT FORMAT 37
2.8 SUMMARY 39

CHAPTER 3. REVIEW OF THE RELATED LITERATURE: FEEDBACK **40**

3.1 PURPOSE AND OUTLINE OF CHAPTER 40
3.2 INTRODUCTION 40
3.3 THE EFFECTS OF FEEDBACK ON STUDENTS 42
3.4 PRE-SEARCH AVAILABILITY 43
3.5 EXTENT OF FEEDBACK 45
3.6 TIMING OF FEEDBACK 48
3.7 SUMMARY 52

CHAPTER 4. REVIEW OF THE RELATED LITERATURE: COMPUTER ASSISTED LEARNING AND ASSESSMENT 53

4.1 PURPOSE AND OUTLINE OF CHAPTER 53
4.2 INTRODUCTION 53
4.3 ILLUSTRATIVE EXAMPLES OF CBA 56
4.3.1 "Question Mark Designer" for Windows at the University of Luton 57
4.3.2 "Dynamic Teaching Solutions" at the University of Western Australia 61
4.3.3 LON-CAPA at Michigan State University and Simon Fraser University 65
4.4 ASSESSMENT AND COMPUTER ANXIETY 76
4.5 THE EQUIVALENCE OF PAPER-AND-PENCIL AND COMPUTER-BASED ASSESSMENT 81
4.6 SUMMARY 87

CHAPTER 5. METHODOLOGY 89

5.1 PURPOSE AND OUTLINE OF CHAPTER 89
5.2 INTRODUCTION 89
5.3 THE ACTION RESEARCH PARADIGM 90
5.4 ACTION RESEARCH AS GENUINE RESEARCH 95
5.5 RIGOUR IN ACTION RESEARCH 100
5.6 RESEARCH GOALS AND ASSOCIATED RESEARCH METHODOLOGIES 106
5.6.1 First research goal 107
5.6.2 Second research goal 108
5.6.3 Third research goal 109
5.6.4 Fourth research goal 110
5.7 RESEARCH METHODS 110
5.7.1 Surveys 111
5.7.2 Interviews with students 114
5.7.3 Focus-group discussions 117
5.7.4 E-mail communications 119
5.7.5 Information collected automatically by OASIS 120
5.7.6 Interviews with instructors 123
5.8 ACTION-RESEARCH CYCLES 125
5.9 ETHICAL CONSIDERATIONS 127
5.10 SUMMARY 129

CHAPTER 6. THE OASIS PROTOTYPE: 2000 - 2002 130

6.1 PURPOSE AND OUTLINE OF CHAPTER 130
6.2 PREHISTORY: 2000 AND 2001 131
6.3 FIRST OASIS ASSIGNMENTS: SEMESTER ONE, 2002 135
6.3.1 Introduction: 'Circuits & Systems' and my role in it 135
6.3.2 Towards the first OASIS assignment 137
6.3.3 The first OASIS assignment 139

6.3.4 Reflections and plans for future action 140
6.3.5 The start of the action research programme 144
6.4 OASIS YEAR-ONE IMPLEMENTATION: SEMESTER TWO, 2002 145
6.4.1 Introduction: 'Electrical Engineering Systems' and my role in it 145
6.4.2 OASIS practice in year one 146
6.4.3 Some statistical explorations 147
6.4.4 Written student feedback 151
6.4.5 Reflections and plans for future action 152
6.5 THE REWRITING OF THE OASIS SOFTWARE 155
6.6 SUMMARY 160

CHAPTER 7. OASIS COMES OF AGE: 2003 162

7.1 PURPOSE AND OUTLINE OF CHAPTER 162
7.2 THE NEW VERSION OF OASIS 163
7.3 THE NEW OASIS SERVER 164
7.4 THE OASIS TRIAL: SEMESTER ONE, 2003 165
7.4.1 Introduction: Changes from 2002 165
7.4.2 The OASIS assignments 166
7.4.3 The OASIS survey, development and implementation 167
7.4.4 Reflections and plans for future action 173
7.5 THE OASIS COMMITTEE 174
7.6 LARGE-CLASS ASSESSMENTS: SEMESTER TWO, 2003 175
7.6.1 Introduction: Changes from 2002 175
7.6.2 The OASIS assignments 176
7.6.3 The OASIS survey 179
7.6.4 Reflections 184
7.7 THE FIRST OASIS STUDENT INTERVIEWS 186
7.8 THE FIRST OASIS STAFF INTERVIEWS 191
7.9 REFLECTIONS AND PLANS FOR FUTURE ACTION 195
7.10 SUMMARY 201

CHAPTER 8. MAINSTREAM OASIS: 2004 203

8.1 PURPOSE AND OUTLINE OF CHAPTER 203
8.2 YEAR-TWO, SEMESTER ONE, 2004 204
8.2.1 Changes from 2003 204
8.2.2 The OASIS assignments 206
8.2.3 Focus-group discussions 207
8.2.4 The OASIS survey 214
8.2.5 Reflections and plans for future action 220
8.3 YEAR-ONE, SEMESTER TWO, 2004 223
8.3.1 Changes from 2003 223

8.3.2 The first OASIS assignment ... 223
8.3.3 OASIS performance upgrades ... 226
8.3.4 The second OASIS assignment ... 230
8.3.5 Student voices: interviews and emails ... 231
8.3.6 The final examination ... 238
8.4 REFLECTIONS AND PLANS FOR FUTURE ACTION ... 238
8.5 SUMMARY ... 246

CHAPTER 9. SPREADING THE RESEARCH NET WIDER: 2005 ... 247

9.1 PURPOSE AND OUTLINE OF CHAPTER ... 247
9.2 YEAR-TWO, SEMESTER ONE, 2005 ... 247
9.2.1 The OASIS assignments ... 248
9.2.2 Focus-group discussion ... 249
9.3 YEAR-ONE, SEMESTER TWO, 2005 ... 253
9.3.1 Changes from 2003 ... 253
9.3.2 The OASIS assignments ... 254
9.3.3 Student voices: interviews and emails ... 256
9.3.4 The final examination ... 263
9.4 OASIS USAGE IN THE DEPARTMENT OF PHYSICS ... 265
9.5 FEEDBACK FROM POSTGRADUATE STUDENTS ... 269
9.6 INTERVIEWS WITH ACADEMIC STAFF ... 273
9.7 REFLECTIONS AND PLANS FOR FUTURE ACTION ... 281
9.8 SUMMARY ... 287

CHAPTER 10. CONCLUSIONS ... 289

10.1 PURPOSE AND OUTLINE OF CHAPTER ... 289
10.2 OVERVIEW OF THE RESEARCH PROGRAMME ... 289
10.3 REVIEW OF THE RESEARCH FINDINGS ... 291
10.3.1 First research goal ... 291
10.3.2 Second research goal ... 293
10.3.3 Third research goal ... 295
10.3.4 Fourth research goal ... 302
10.4 GUIDELINES FOR SUCCESSFUL IMPLEMENTATION OF OASIS ... 302
10.5 QUALITY / VALIDITY OF THE RESEARCH ... 304
10.5.1 Outcome validity ... 304
10.5.2 Process validity ... 305
10.5.3 Democratic validity ... 306
10.5.4 Catalytic validity ... 307
10.5.5 Dialogic validity ... 307
10.6 IMPLICATIONS AND FUTURE DIRECTIONS ... 308
10.7 IN CONCLUSION ... 310

REFERENCES **311**

APPENDICES **324**

APPENDIX ONE: OASIS SURVEY FORM 324
APPENDIX TWO: INFORMATION SHEETS AND PERMISSION FORMS 326
APPENDIX THREE: OASIS PRACTICE; SCREENSHOTS OF REPRESENTATIVE STUDENT PAGES 331
APPENDIX FOUR: OASIS ASSESSMENT; SCREENSHOTS OF REPRESENTATIVE STUDENT PAGES .. 338
APPENDIX FIVE: OASIS STAFF-ONLY PAGES, REPRESENTATIVE SCREENSHOTS 339
APPENDIX SIX: OASIS WISH LIST 2002; STAFF ACCESS TO DATABASE 349
APPENDIX SEVEN: SPSS ANALYSIS OF OASIS PRACTICE DATA 353
APPENDIX EIGHT: SURVEY ANALYSIS, 'CIRCUITS & SYSTEMS', 2004 355
APPENDIX NINE: PUBLICATIONS LIST ASSOCIATED WITH OASIS 356
APPENDIX TEN: ABBREVIATIONS 357

Chapter 1. Introduction

1.1 Purpose and outline of chapter

This introductory chapter provides background and context for the research and presents an outline of the research goals and questions. It is argued, in Section 1.2, that increasing student numbers and diversity, and therefore increasing instructor workload, together with the increasing affordability and capability of computer software and hardware, provide compelling motivation for the adoption of computer-assisted learning (CAL) and computer-based assessment (CBA) and a move away from paper-and-pencil assessment (PPA). The implementation of an early version of OASIS (Online Assessment System with Integrated Study), the software package that is the basis for my research, is outlined in Section 1.3. This section also discusses the decision taken by Department of Electrical and Computer Engineering (ECE) at the University of Auckland (UoA) to develop OASIS further and to investigate its implementation. Note: abbreviations are listed in Appendix 10. A broader rationale for the research, in terms of its contribution to the global research community, is provided in Section 1.4. The research goals and questions are presented in Section 1.5, while the research design is briefly described and justified in Section 1.6. The significance of the research is outlined in Section 1.7, where it is noted that OASIS possesses features that distinguish it from a number of other similar software packages. Some limitations of the study, particularly with respect to the generalisation of the results, are noted in Section 1.8. Finally, a chapter summary is provided in Section 1.9.

1.2 Background

Individuals pursuing an academic career will be all too aware of increasing student numbers and increasing staff workloads in higher education. My own experiences certainly bear testimony to this. I graduated from UoA in 1972, then an institution of some 8000 students. When I returned as a member of the academic staff in 2002 the student numbers were approaching 30 000. My first- and second-year classes in ECE ranged in size from about 220 to 350 students. Even some year-four classes contained more than 160 students. In my time away from university I taught physics

at Rangitoto College. In 1974 there were just 12 students studying year-13 physics. In my final years of teaching the number of students studying year-13 physics had increased to about 120 and class sizes had doubled to 24 students. However, the number of classes taught by teachers had not decreased.

The anecdotal evidence above is backed by the more systematic studies of others. A recent study conducted as part of the Faculty's IPENZ (Institution of Professional Engineers New Zealand) accreditation process showed that the numbers of students enrolled at UoA's School of Engineering in general and in ECE in particular have increased significantly. In 1998 the School had 1344 undergraduates and this grew to 2010 undergraduates in 2003, a 50% increase in a five year period. Further afield, in the United Kingdom, in the mid-sixties the median lecture audience size was found to be 19 (Hale, 1964), while another survey revealed that the average number of students in a discussion group was just 4.1 (Robbins, 1963). These figures are scarcely credible in today's world. Biggs (1999) reported that over a 10-year period the percentage of school leavers in higher education increased from 15% to 40% which has resulted in increased class sizes because funding has not increased in proportion. In the UK, per student funding in higher education decreased by over 50% in real terms from 1985 to 1990 and similarly increased student numbers combined with decreased funding can be observed in most other European countries (Dill & Sporn, 1995) with the US also unable to avoid a period of constrained educational resources (Sporn, 1999).

The demographics of most countries are currently in a state of rapid change: more and varied ethnic backgrounds, age groups and minorities are being represented in society in general and therefore in higher education in particular. These demographic changes are actually even more pronounced in higher education than in society in general because the fraction of the population now going on to higher education is increasing dramatically. This increasingly diverse student population would best be met with more individualised treatment and a greater variety of educational programmes. However, the increasing class sizes brought about by decreased funding tend to force the opposite: a more standardised approach to education (Biggs, 1999; Fernandez, John, & Netherwood, 2001; Sporn, 1999; Stephens, Bull, & Wade, 1998).

One side-effect of increasing student numbers is the effect it has on the *type* of work carried out by academics. As classes become larger, the assessment load increases too, and academics find themselves spending a greater fraction of their time on assessment and a smaller fraction on preparing and delivering lectures. Not only is this likely to be less satisfying for academics, but it is likely to impinge negatively on the quality of teaching and learning, already under threat from increased class sizes. One recent analysis carried out by Excell (2000) in an electrical engineering context showed that in classes of 100 or more students, the lecturer may well spend less time on lecturing, lecture preparation, tutorials, etc. than on the final assessment alone. Excell gives figures of approximately 100 hours for the final assessment (range: 60 to 170 hours) and approximately 75 hours for lecturing, lecture preparation, tutorials, etc. Assignments, tests and laboratory reports would add further to the assessment load, making an even greater imbalance. Larger classes would also produce relatively more hours for assessment. Table 1.1 shows the planned staffing time allocations for some representative courses in ECE in 2003. The actual student numbers proved to be somewhat greater than those in the table and time allocations were increased accordingly, but the data show clear recognition of the fact that the bulk of academic time in large courses is devoted to assessment. For courses of 100 students, about half the time is spent on assessment and this fraction increases with class size.

Table 1.1
Time allocations for some representative courses in ECE, 2003

Course	Year level	Number of students	Hours allocated for lecture preparation and delivery	Hours allocated for assessment
ElectEng 101	1	525	227	525
ElectEng 303	3	230	206	345
ElectEng 302	3	90	158	135
CompSys 402	4	35	158	53

With the large class sizes now prevalent and much of the increased workload being in the area of assessment, academics are likely to cut back in this area in order to

hold workloads at manageable levels. There are at least three distinct approaches available for academics wishing to make these cuts. Perhaps the simplest approach is to reduce the number of assessments. Since the final examination is unlikely to be abolished, it is likely that this approach will entail reduced formative assessment. The role of formative assessment in motivating and enhancing student learning will be discussed in detail later, suffice it to note here that its role is regarded as extremely important and its reduction is likely to have significant adverse effects on student learning (Black & Wiliam, 1998a; Crooks, 1988). Indeed, one highly respected commentator has stated that "formative assessment… is at the heart of effective teaching" (Black & Wiliam, 1998b, p. 140).

The second approach entails maintaining the number of assessments, for the reasons hinted at above, but finding a way to produce and/or mark them more easily. Multiple choice and other test formats that lend themselves to rapid marking are becoming more popular. For example, at the University of Luton there has been a significant shift to final examinations that are largely multiple-choice: a typical final examination is one hour long, contains about 60 questions, and is taken by a class of 140 students (Zakrzewski & Bull, 1998). Academic staff have reacted very favourably, commenting in particular on the reduction in marking effort, the increase in marking consistency, and the fast availability of results. While some commentators have noted with some disquiet that an overemphasis on multiple choice testing appears to be developing, particularly in large classes (Paxton, 2000), others have made the point that one type of testing format is not inherently superior to another: rather it is important that the assessment format is appropriate for testing the particular skills and content that are to be evaluated (Crooks, 1988).

A third approach to managing assessment workload is to at least partly automate the process of marking. This approach is often combined with a shift to multiple-choice or similar test formats: clearly computers can mark multiple-choice easily and reliably but are much less well suited to essay marking. There are many universities where computers are used to reduce the assessment workload (Batchelor & Jungic, 2004; Scott & Stone, 1999; Sly & Rennie, 1999; Tartaglia & Tresso, 2002; Thoennessen & Harrison, 1996), and some of these will be looked at in more detail in Chapter 4.

This third approach is becoming more attractive: in spite of tighter budgets, computers are becoming more prevalent and more useful as their cost drops and their power increases. In 1974, my first year of teaching, Rangitoto College was one of the few schools to own a computer, a Digital PDP8. It cost more than three times my annual salary and, when the language was loaded, there was 0.8k of RAM left to run programmes. Now a starting teacher's annual salary could purchase a class set of 30 computers, and they would have computing power unimagined in 1974. In my last year at UoA, I regularly booked the University's mainframe computer for an eight hour programming session for my research. The following year that computer was replaced with one 80 times more powerful. My eight-hour programming sessions could have been completed on the new computer in six minutes!

Aizcorbe, Corrado, and Doms (2000) tracked desktop and laptop prices over a five-year period and observed that prices fell by almost 30% annually. Nordhaus (2002) noted that the real power of computing per dollar increased by 70% to 80% per annum for the period 1980 to 2000. With the significant and continued growth in what is achievable by computers given a reasonable budget, it is not surprising that some commentators have suggested that computers can provide a partial solution to the problem of increased student numbers and diversity (Reinhardt, 1995).

Several computer-based tools such as WebCT (2004), Questionmark (2004) and Blackboard (2004) have been developed in order to assist instructors in the delivery and assessment of courses. WebCT was trialled in year-one courses by the Faculty of Engineering at the University of Wollongong, Australia and found to be a most useful tool, to the extent that increased usage was planned for the future (Baafi & Boyd, 2001). Rather than purchase an off-the-shelf solution, UoA developed Cecil (2006). Cecil performs its main role as a course management tool very well, and has a number of powerful and useful features. For example, Cecil's course database is updated each night through its linkage to nDeva (2006), a software package that handles the University's enrolment program. Thus instructors can rely on the class roll being correct. Further, instructors are automatically notified by email of any changes in class roll. It is also a simple matter for instructors to send every student enrolled in any given course an email and all such emails are automatically logged.

This is a much more reliable way to disseminate critical information than the alternative, an announcement in a lecture. Cecil also provides instructors with the details for each student, including an ID photograph.

Cecil also has some assessment capabilities. However, ECE has found these to be somewhat limited in terms of available question format and ability to handle large groups of students quickly and reliably. Cecil is not alone in having limited question formats; even a cursory glance at the literature would suggest that most computerized assessment systems are similar, often being restricted to multiple-choice questions (Excell, 2000; Nafalski & Samaan, 1997; Tartaglia & Tresso, 2002), or perhaps multiple-choice questions and variations on that theme, such as choosing an answer from a drop-down list.

1.3 Context of study

My first acquaintance with UoA was as a student of physics and mathematics. After graduating with degrees in these subjects, I then proceeded to teach them in high school for 27 years. During this time I also pursued some part-time university study, firstly obtaining a bachelor's degree in philosophy from UoA and then, more recently, taking some doctoral-level papers in science education from Curtin University. It was during this latter part-time study that I was offered a job at UoA in the Department of Electrical and Electronic Engineering (at the start of 2004 the department changed its name to the Department of Electrical and Computer Engineering). As a result of this, I joined the engineering faculty's academic staff at the start of 2002 with very little engineering in my background, considerable practical experience in education, and looking for a new area of research that could be pursued for my doctorate. By good fortune, such an area rapidly made itself known to me.

During 2000, the department had developed OASIS (Online Assessment System with Integrated Study), a software package described as a web-based tutorial and examination system (Bigdeli, Boys, Calverley, & Coghill, 2001). OASIS consisted of a question bank and a database that recorded student activities. Students could log on to OASIS where-ever they had internet access, select a course, then a topic area, and

then a question to attempt. On submitting their answer to the question, the students would be told whether they were right or wrong, and informed of the correct answer. Each time an answer was submitted, the database would store the student's ID, the answer submitted, the correct answer, and the time at which the answer submission occurred. Virtually all the questions in the question bank involved the solution of electrical circuits. The answers to these questions were numerical, and student answers were deemed to be correct if they were within 1% of the correct answer. If a student chose to repeat a particular question, he or she would be provided with the same circuit configuration but with different numerical values. This was a useful facility: frequently students wish to practise a particular circuit configuration several times in order to feel that they have achieved mastery.

OASIS was available in two different modes. Students initially met OASIS in practice mode. In this mode, they could log on to OASIS, practise questions, and improve their circuit-solving skills. Students received instant feedback about their efforts and the repeatability feature of OASIS gave it a potential advantage over problems set from the textbook. Practice mode also gave students the opportunity to familiarise themselves with the OASIS environment.

OASIS was also used in test mode. This required all students in a particular course to log on to OASIS at approximately the same time in a supervised environment. These students would then be given an hour to complete a test assembled by their instructor from the set of questions they had previously been able to practise. This, of course, gave the students a great incentive to practise questions prior to the test. However, since each question had about 200 different numerical versions, students could not memorise answers to questions. Instead, in order to be successful in the test, students needed to be able to select suitable methods for solving the given circuit configurations. When students submitted their answers to the test questions, they were given instant feedback. Instructors also received comprehensive feedback. The computerized marking performed by OASIS resulted in test scores being available about two weeks earlier than they would have done had the marking been carried out in the traditional fashion.

In 2000 and 2001, OASIS was used by a few academics of somewhat pioneering spirit to provide practice and test opportunities in a few second and third-year courses. These academics were, in general, most enthusiastic about the potential of OASIS. It was seen as providing a way of reducing the crippling workload associated with assessment. It was also judged to have significantly lifted student skill levels in the important area of circuit analysis. There was some scepticism in the Department however, with a few staff wondering whether OASIS was really promoting problem-solving skills or simply, through repetition, enhancing the ability of students to solve circuits they had previously met anyway. Some preliminary research did indicate that student performance in traditional written tests had noticeably lifted as a result of providing students with practice opportunities on OASIS prior to the tests (Bigdeli, Boys, Calverley, & Coghill, 2001). A survey of student opinion indicated that a clear majority (73%) also felt that OASIS had helped them improve their skills (Bigdeli, Boys, Calverley, & Coghill, 2001).

OASIS had been created, not in opposition to Cecil, but rather to be complementary. Cecil was used as a course management device, recording student results, providing students with electronic materials, e-mailing announcements to students and so on. In this respect, instructors found it to be a powerful and useful tool, and were happy to employ it. However, Cecil's assessment facilities were seen as a poorly-featured add-on to its main functions. The multiple-choice testing it provided was seen as much too limited. Furthermore, it was felt that being able to repeat problems over and over again if necessary with different numerical values each time was highly desirable. Some instructors had actually used Cecil to provide practice and testing opportunities for their students. Unfortunately, their experiences had not always been favourable. For example, one of my colleagues ran a test for a second-year class of 160 students. The class was divided into two sessions, so that 80 students were tested in each session. The test was in two parts, the first was a set of multiple-choice questions delivered by Cecil, while the second part was a set of questions delivered by OASIS. This further division resulted in only 40 students using Cecil at one time. In spite of this, the group of students using Cecil found that the system was responding so slowly to their inputs that that half of the test had to be discounted. No such speed problems had been encountered with OASIS even when 100 students were simultaneously being tested by it. Classes of 200 students had also been able to use

OASIS consistently without any delays being reported. The preference of instructors was to use Cecil as a course management tool and OASIS for skills practice and assessment. This combination proved to be most successful, being well received by both students and instructors (Hussmann & Bigdeli, 2002; Hussmann & Smaill, 2003).

The above outlines the situation as I found it on my arrival in the Department at the start of 2002. Sufficient problems had been written to allow OASIS to partially cover a few second- and third-year courses. Producing questions for OASIS was actually a non-trivial exercise due to the fact that each question required about 200 different numerical versions. These numerical versions were generated before the question was made available to students and stored in the question database. When the question was requested by a student, one set of data would be selected from the database, this set containing the numerical values needed for the question as well as the answers corresponding to those numerical values. This feature made OASIS very quick: marking a student's answer involved comparison rather than calculation. All calculations had been carried out before the student even requested the problem. On the other hand, this feature also made it more difficult to write questions for OASIS: a program such as Matlab had to be used to generate the array of values needed for the question. Typically, a question would be designed by the instructor while a senior student possessing the requisite experience would generate any graphics required as well as the set of numerical values and answers for the question. The question itself, together with graphics and number-set, would then be loaded into OASIS. Any extension of the use of OASIS would obviously require more questions. This would require a time commitment from instructors to write the questions, funding from the Department to pay the senior students to encode the questions, and further time commitment from instructors to check that the questions as delivered by OASIS were functioning correctly.

Another constraint on the expansion of OASIS was the software itself. The software had been written, to some extent, as a prototype. It was becoming unable to support the increasing demands of more and larger classes. Expert staff time was also needed to maintain and configure it. The software had developed in a somewhat haphazard fashion and this made it difficult to extend the software to handle some new question

types that were being requested by instructors. As more classes and more questions were added to the system, navigation became awkward. Furthermore it became difficult for instructors to extract information and results for individual classes. In fact, the development of the prototype had reached the point where, if further development was to proceed, some very significant modification or complete rewriting would be required.

Thus, in early 2002, the time at which I joined the Department, the point had been reached where an important decision needed to be made concerning the future of OASIS. If its implementation was to be more widespread, significant staff time and department funding would need to be allocated to it. Clearly, this allocation of resources would need justification. If it was found that staff and students were both enthusiastic about its use, then that would be at least partial justification for undertaking to develop and expand OASIS in a major way. However, the most important issue of all, student learning, also needed to be recognised. It would be critical to gather information on the learning of students and attempt to determine whether the implementation of OASIS was actually enhancing that learning.

The department already had a member of the academic staff who was allocated one day a week for work involving OASIS. His brief was to maintain the prototype version of OASIS and to develop a new version. Therefore, my role in this area was to develop a list of requirements rather than to write any lines of code. As a member of the teaching staff who wished to use OASIS with his classes, I would also be involved in writing questions and implementing the usage of OASIS. Further, my interest in educational research meant that I was keen to study the implementation of OASIS, determine staff and student reaction to it, and establish to what extent student learning had been enhanced through its use. For reasons that will be discussed later, action research was chosen as the research methodology.

1.4 Study rationale

First, it must be noted that the Department had its own rationale for conducting this piece of research. In order to take OASIS from the prototype stage and implement it in a more widespread fashion, a considerable ongoing investment of money and

academic time would need to be made. It was important to be able to justify this in terms of learning outcomes and in terms of instructor and student satisfaction. The research could also be used to guide the way in which OASIS was developed and implemented in courses.

Second, while any well-conducted piece of research can be justified in the sense that it adds to the body of research knowledge, it is particularly easy to justify further research in the area of CAL. This is because the environment is continually changing. As computers become more powerful, they are able to perform tasks that they could not have previously performed. Further, as computers become more affordable, their use becomes more widespread, resulting in both instructors and students becoming more familiar with them and being able to use them for an increasing range of tasks without "computer anxiety". This increasing familiarity with computers can quickly invalidate an earlier research result, particularly one that found that students did not benefit when a certain learning approach was implemented by computer. It may have been that when the study was conducted, for example, a significant number of the students were very new to computer use, or that a significant number had very poor access to the internet. In this situation, a more recent study could well produce a quite different outcome. The very newness of CAL may mean that research findings are still to be determined on some issues. For example Bonham, Beichner, and Deardorff (2001) reported that students presenting homework online experience no significant differences in learning gains when compared to students completing homework in the traditional pen and paper way. They also found no significant differences in examination solutions of a group of students presenting homework online and another group presenting homework traditionally. However, other researchers (E. Kashy, Thoennessen, Tsai, Davis, & Wolfe, 1997) have found that online homework affected student learning in a significant and positive way and that it promoted expert problem-solving techniques. Further research in this area is needed to resolve this and other differences.

Third, this piece of research targets two areas of critical importance in the areas of teaching and learning. The first of these is the use of computers in education. This is of critical importance because the use of computers is very widespread in education and in the future it will become even more so. Thus, it is important to gather up-to-

date information on the most effective way to implement computers in education. The second is assessment; also of critical importance because, in a very tangible way, the assessment system creates the 'hidden curriculum'. It determines student activities, affecting the amount of effort they make and how they allocate their efforts (Gibbs, 1999; Innis, 1996; Sambell, Sambell, & Sexton, 1999). OASIS is currently used to provide students with practice opportunities, assignments and in-semester tests, all with prompt feedback. Thus its primary role is formative assessment rather than summative assessment. In a landmark study of the available research literature, Black and Wiliam (1998a) put an unassailable case for the significant positive learning gains to be made by the effective implementation of formative assessment:

> there is a body of firm evidence that formative assessment is an essential component of classroom work and that its development can raise standards of achievement. We know of no other way of raising standards for which such a strong prima facie case can be made. (p. 148)

With assessment, and particularly formative assessment, playing such a crucial role in learning, it is easy to justify research that aims to enhance student learning through studying the implementation of a software package that delivers formative assessment opportunities to students.

1.5 Research goals and questions

The overall aim of this research was *to develop, implement and validate a Web-based software package that, through providing practice and assessment opportunities, improves student learning and reduces marking and related mundane aspects of instructor workload*. This overall aim was translated into four research goals. For each goal, research questions were written to help focus on the goals and determine the extent to which they had been achieved. The following subsections present each of the goals in turn, together with their associated questions.

The methodology was action research, which, owing to its spiral nature, enabled each research goal and its associated research questions to be visited more than once. The responsive nature of action research also left open the possibility that, as the research

proceeded, the initial research goals and questions could change somewhat or lead to more focused goals and questions.

1.5.1 First research goal

To produce a software package (OASIS) that meets certain criteria (e.g. Web-based, secure, robust, scalable, easy to use, provides prompt feedback, logs student activities and results) and can be used for student practice, assignments and tests in electrical engineering.

Associated research questions

1a. What criteria are appropriate for OASIS to meet?

1b. Does OASIS meet the criteria?

1c. What student activities and results are appropriate for OASIS to log?

1d. Does OASIS log the required student activities and results?

1.5.2 Second research goal

To incorporate features in OASIS that encourage staff to adopt it for their use with courses they teach. In particular, use of OASIS should reduce staff workload. This goal also includes the longer-term goal of extending the use of OASIS beyond the department, faculty and University.

Associated research questions

2a. Which instructors are using the package and in what courses?

2b. Why did these instructors decide to use OASIS?

2c. What features of OASIS are appreciated by instructors and what features need changing, deleting or adding?

2d. What effect is the use of OASIS having on instructor workload?

1.5.3 Third research goal

To implement the use of OASIS in such a way that student learning is enhanced.

Associated research questions

3a. How do students use OASIS?

3b. What features of OASIS are appreciated by students and what features need changing, deleting or adding?

3c. What student OASIS usage patterns correlate with increased student learning?

3d. How do instructors implement the use of OASIS?

3e. What instructor OASIS implementation patterns correlate with increased student learning?

1.5.4 Fourth research goal

To identify best practice strategies for the implementation and use of OASIS.

Associated research questions

4a. How should instructors implement and use OASIS to best promote student learning?

4b. How should students use OASIS to best promote their learning?

While the above research questions focus on the development of the software package itself, it should be understood that in parallel with this development is the writing of questions to be delivered by OASIS to the students. Clearly, the success or failure of OASIS as a learning tool depends not only on the software itself but also on the questions it delivers to students. Of course, OASIS can help instructors produce effective questions. The timely feedback instructors receive about student performance on the questions may well lead to questions being modified, supported by further questions, or being deleted. There are several instructors involved in writing questions because OASIS is used in several different courses.

1.6 Brief description of the research design

If this research had had the single aim of establishing whether the use of OASIS enhanced student learning, then a different research design to the one actually employed might have been considered. It would have been very tempting to simply divide the students in a course into two equivalent groups, providing one group with access to OASIS while denying the other group access. Comparing the performances of the two groups in tests and the final examination would then have given a useful quantitative indication of the worth of OASIS. However, there were two arguments against this approach. First, when I began my research, a prototype version of OASIS had already been used in a few courses. In spite of the fact that its use had really only been sporadic, and its coverage of course content anything but comprehensive, there was already a substantial body of anecdotal evidence that suggested that student learning was being significantly enhanced by the use of OASIS. Therefore, it would

have been difficult to arbitrarily deny OASIS access to half the students on a course. Any instructor taking this step would have felt guilty about the strong possibility of disadvantaging half of the students, and the students themselves certainly would have felt disadvantaged and would have complained vociferously!

The second argument against the research approach outlined above was a natural consequence of my double role as both instructor and researcher. As instructor, I aim to maximise the learning of my students. Therefore, I cannot simply adopt a research design that shows a certain piece of software enhances student learning. Rather, my aim must be to investigate the use of the software and, if possible, modify both the software itself and its usage so that the learning of my students is maximised.

The requirement for a responsive approach that included not only research but also encouraged improvements to teaching and learning practices led me naturally to adopt the methodology of action research. Elliott (1991) suggests that action research might be defined as "the study of a social situation with a view to improving the quality of action within it" (p. 69). This was certainly in line with my intentions. Carr and Kemmis (1986) see action research as proceeding through a spiral of cycles of planning, acting, observing and reflecting. For my action research, each cycle would naturally take one semester, since in each semester I would teach a complete course, including all assessment. Planning would take place before the semester, acting and observing during the semester, and reflecting after the semester. This reflection would inform the planning for the next cycle. My teaching schedule always provided me with an adequate non-teaching period between semesters. Therefore, there was sufficient time between cycles for activities related to observation (such as analysing interviews and data collected by OASIS), reflecting and planning.

By choosing action research rather than the two-groups approach outlined in the first paragraph, I had actually made the task harder. First, I was committing to two activities rather than one: improving my courses and carrying out research. Second, I was abandoning much of the security of the statistical analysis that might have resulted from a more quantitative approach. Without the control and manipulation of the two-groups approach, it would now be particularly important to convincingly establish the validity and reliability of a piece of research that some would view as

overly qualitative. Working in the Faculty of Engineering, I was all too aware of the high esteem in which quantitative research can be held. In fact, considerable quantitative data would be analysed in the course of this study. OASIS itself would collect data every time a student used the programme. Test and examination results would also be analysed. Finally, questionnaires would also yield quantitative as well as qualitative data. Further qualitative data would come from observations, informal discussions, and more formal interviews with both students and staff. Most, but not all, of the interviews would be carried out by myself. Additional triangulation would be achieved by continuing the study through several cycles over several semesters.

It should be noted that OASIS is used in several different courses, and I am not involved with all of these courses. Some of the courses that have recently adopted OASIS are not in my Department or Faculty. In order to maintain sufficient depth in my research, my attention would be given preferentially to the courses that I lecture: Circuits & Systems, a year-two first-semester class of about 200 students, and Electrical Engineering Systems, a year-one second-semester class of about 550 students. Thus, student interviews would be restricted to these classes, although staff interviews would be representative of all past and present users of OASIS.

More details of the research methodology are provided in Chapter 5.

1.7 Significance

Locally, this study claims significance because it aims to validate the use of OASIS within the Department by establishing the extent to which it improves student learning and assists instructor activities. Further, in order to provide the maximum benefit, it sets out to improve OASIS itself and the way in which it is implemented in courses. The feedback that OASIS provides about student practice and assessment activities to instructors should enable them to produce more effective teaching programmes. For example, areas of difficulty for students can be identified and instructors can go through these key areas carefully during lectures or tutorials.

Globally, it is hoped that this study also will be seen as having some significance. It has been said that change does not occur as the result of a single study but rather as

the result of a steady accumulation of research and experience. This statement is particularly applicable to areas such as education research that are frequently the subject of qualitative rather than quantitative studies. The present study adds its voice to the accumulation of knowledge in the important areas of CBA and CAL. Given the rapid ongoing changes in any areas that involve computers, research from alternative theoretical positions is not only welcome but in fact necessary in such areas.

This study also targets some areas that appear to be under-represented in the research literature. In some cases this may be because OASIS offers some features or is used in some ways that are seen with very few other packages. The present research does not make much use of OASIS's ability to deliver multiple-choice questions, focusing instead on other question types. However, multiple-choice questions would seem to be the most popular question type delivered by CBA packages (Bull & Collins, 2002), and, perhaps as a consequence, the bulk of the related research literature focuses on their delivery. Also, a number of packages, including some well-established ones, do not offer students access to practice questions throughout the course. Further, the assignments delivered by OASIS, together with the practices surrounding them, appear to constitute a unique package. I have not found any examples of institutions whose practices in this area replicate ours.

The area of formative assessment is also under-represented in the research literature. This may be because, in general, formative assessment is under-utilised in practice while too high an emphasis is placed on summative assessment (Black & Wiliam, 1998a; Crooks, 1988; Hunt & Pellegrino, 2002). The present study has a strong emphasis on formative assessment. The practice questions that students use throughout their courses offer prompt feedback and are clearly formative rather than summative in nature. Even the assignments, which are summative in the sense that performance on them contributes to the final course mark, have a formative component by virtue of their timing: they are held throughout the semester and the feedback that students receive from them is delivered sufficiently early to enable study habits to be adjusted appropriately. For example, the first OASIS assignment generally takes place only two weeks after the start of the semester.

Finally, the data collected by OASIS on student performance, as well as being useful to instructors, provides an extremely valuable research resource. The information gathered is comprehensive in nature and affords a rare insight into the actual work patterns of students. Relatively few researchers appear to have mined such data. This may be because the software packages under their study did not collect such data or did not provide ongoing practice opportunities for students.

1.8 Limitations

One limitation of this study is that it is based on the implementation of a software package that is used in only one institution. For this reason it is difficult to generalise from its findings. This limitation is shared by a number of other studies; in fact the number of institutions that have chosen to produce their own software packages for computer-assisted assessment is as great as the number that have purchased off-the-shelf solutions (Rothberg, Lamb, & Wallace, 2001). Even the most popular off-the-shelf solutions do not enjoy widespread popularity on an international scale.

Another limitation faced by this study was that all investigation took place in courses that were being run in the normal sort of way. All decisions concerning practice and assessment opportunities were made from the perspective of an instructor concerned with student learning rather than from the perspective of a researcher. This requirement that the learning environment be as natural as possible precluded manipulating it to suit research ends. For example, it would have gone totally against the philosophy of this study to deny one group of students access to OASIS while allowing another group to use it. This sort of research has been conducted elsewhere (Pascarella, 2002). In case the completion of questionnaires is seen as unnatural, it should be noted that engineering students routinely complete several questionnaires each semester about learning and teaching. Data collection had to be as unobtrusive as possible, and OASIS was the ideal tool for this, being able to collect considerable data automatically without students being aware of it (although they were all advised at the start of the course that it would be happening).

Finally, as for any piece of research of this kind, the exact circumstances under which it was conducted are unique. Therefore its findings may not be applicable to

other situations that appear similar, and investigators working in these may even find results that differ from those found in this study.

1.9 Summary

Growing class sizes are producing challenging instructor workloads, particularly in the area of assessment. The increasing affordability and sophistication of computer software and hardware offer a solution to this problem. The software package OASIS was produced to provide formative and summative assessment opportunities for students in ECE. The overall aim of the research described in this thesis was to develop and implement OASIS so as to demonstrably improve student learning and reduce instructor workload. An action-research methodology was chosen as being the most appropriate for achieving the goals in an ethical fashion. This methodology did impose some limits on the sort of research that could be carried out; in particular, control groups could not be used. The research was also limited by the fact that it took place in a single institution and involved a small range of subject areas. However, the research still has significance because it adds to the body of research in a rapidly-changing area and because OASIS offers some features not often found in other similar software packages.

Chapter 2. Review of the Related Literature: Assessment

2.1 Purpose and outline of chapter

This chapter reviews the literature on assessment in general, while Chapters 3 and 4 focus on the related areas of feedback and CBA. Some statements of definition and purpose for assessment are examined in Section 2.2. Summative and formative assessment are then discussed in Sections 2.3 and 2.4 respectively. The point is made that, while summative assessment is generally assigned supreme importance, formative assessment is better positioned to significantly improve student learning. The profound effect assessment has on students is more closely examined in Section 2.5, where it is noted that assessment informally defines the curriculum, largely determining the extent and area of student efforts. The effects of assessment frequency and format are explored in Sections 2.6 and 2.7 respectively. Finally, a summary of the chapter is presented in Section 2.8.

2.2 Introduction

Lambert and Lines (2000) described assessment as "the process of gathering, interpreting, recording and using information about pupils' responses to educational tasks" (p. 4). Thus, assessment could involve such activities as a teacher observing a student at work in the laboratory as well as a student completing a multiple-choice test. The single sentence quoted above leaves open the question of how the information collected is in fact used. Bloom, Hastings, and Madaus (1971) take it a step further when they state that assessment "is the systematic collection of evidence to determine whether in fact certain changes are taking place in the learners as well as to determine the amount or degree of change in individual students" (p. 8). Linn and Gronlund (2000) echo the above notion that assessment involves gathering information about student learning achievement and put this in the context of learning goals and judgements concerning whether the learning goals have been met:

> Assessment of student learning requires the use of a number of techniques for measuring achievement. But assessment is more than a collection of techniques. It is a systematic process that plays a significant role in effective teaching. It begins with the identification of learning goals and ends with a judgment concerning how well those goals have been attained. (p. 29)

In reality, few accounts of assessment in the literature end at the point where the judgment is made of the learner's achievement. The purpose for which the judgment is made is also regarded as important and in fact the type of assessment being carried out is defined in terms of this purpose. Bell and Cowie (2001) note that assessment has multiple purposes, and that "These multiple purposes can include auditing of schools, national monitoring, school leaver documentation, awarding of national qualifications, appraisal of teachers, curriculum evaluation and the improvement of teaching and learning" (p. 2). Two main purposes are discerned. When the purpose of the assessment is to "provide feedback to teachers and pupils about progress in order to support future learning" (Lambert & Lines, 2000, p. 4) then the assessment is regarded as formative assessment. When the purpose of the assessment is to "provide information about the level of pupils' achievements at points during and at the end of school" (Lambert & Lines, 2000, p. 4) then the assessment is regarded as summative assessment. The literature specifically concerning these two types of assessment is reviewed in the following two sections of this chapter.

2.3 Summative Assessment

In line with the last quote in the previous section, Bloom, Hastings and Madaus (1971) also see one primary goal of summative assessment as being the grading or certifying of students. However, to this they add two other primary goals, namely judging the effectiveness of the teacher, and judging the curriculum:

> Perhaps the essential characteristic of summative evaluation is that a judgment is made about the student, teacher, or curriculum with regard to the effectiveness of teaching or instruction, after the learning or instruction has taken place. It is this act of judgement which produces so much anxiety and defensiveness in students, teachers, and curriculum makers. (p. 117)

In defining summative assessment, Linn and Gronlund (2000) also mention these two other goals, while rating them as somewhat less important than the goal of grading or certifying students:

> Summative assessment typically comes at the end of a course (or unit) of instruction. It is designed to determine the extent to which the instructional goals have been achieved and is used primarily for assigning course grades or for certifying student mastery of the intended learning outcomes.... Although the main purpose of summative assessment is grading, or the certification of student achievement, it also provides information for judging the

> appropriateness of the course objectives and the effectiveness of the instruction. (p. 41-2)

Summative assessment is typically high stakes assessment. The results of such assessment may well determine a student's future study and career prospects. Although summative assessments generally produce less anxiety in the instructors and curriculum-makers associated with a course, they too can suffer unpleasant consequences if the students' achievements overall are far removed from the norm. It is therefore understandable that such assessments often produce anxiety. This anxiety is particularly pronounced when the assessment comes at the end of the course, for in this situation there is no chance for a poor result to be remedied. Student anxiety may also be increased when students feel that they are unfamiliar with the assessment task or the circumstances in which it is presented (Crooks, 1988).

Currently, in most courses, summative assessment consists of more than one final examination. Tests, assignments, laboratory exercises, and the like may also be involved. This has a number of advantages. First, it somewhat reduces the importance of the final examination and therefore reduces the level of anxiety associated with it. Second, it spreads the summative assessment over a larger period of time and perhaps also over a range of different assessment activities, therefore offering a more accurate judgment of student achievement (Black, 1998; Sly & Rennie, 1999). Third, student activity will be spread somewhat more evenly through the course, rather than just concentrated into a short time period prior to the final examination (McDowell, 1995; Miller & Parlett, 1974; Snyder, 1971). Fourth, the experience of taking part in the earlier assessments and the feedback gained from them may help the students to lift their achievement in the final examination (Sly & Rennie, 1999). For this reason the border between formative and summative assessment is somewhat blurred (Light & Cox, 2001). Even tests or examinations that occur at the end of a unit can be regarded as formative where there is a related following unit (Linn & Gronlund, 2000). Some of the above issues will be further explored later in this chapter.

After an extensive literature review, Crooks (1988) concluded that the emphasis on and importance attached to summative assessment can have a number of undesirable consequences such as

> reduction of intrinsic motivation, debilitating evaluation anxiety, ability attributions for success and failure that undermine student effort, lowered self-efficacy for learning in the weaker students, reduced use and effectiveness of feedback to improve learning, and poorer social relationships among the students. (p. 468)

Black and Wiliam (1998a), after a substantial review of the literature, similarly concluded that performance goals lead to lower motivation and achievement outcomes than do learning goals. Therefore, in some learning situations, poorly managed summative assessment regimes can actually be counterproductive.

The emphasis and reliance placed on summative assessment is also unfortunate given that the results of some assessments are not particularly accurate. This is not necessarily the fault of the assessment item itself. For example, the type of assessment used at university level has the potential to influence student grades (Brown & Knight, 1994; Sly & Rennie, 1999). Even the timing of an assessment may radically alter the outcome. One investigation (Black, 1963) involved two 'parallel forms' of a test, using the same types of questions and covering the same learning aims. The tests were given on two occasions only a short time apart and the two sets of results were compared. Amazingly, half of those who failed one paper passed the other, and vice versa.

2.4 Formative assessment

By contrast with summative assessment, formative assessment is not used to grade students. Rather, its purpose is to provide feedback to students and teachers about progress to date so that students can improve their learning and teachers can improve their instruction (Bell & Cowie, 2001).

> formative assessment is used to monitor learning progress during instruction. Its purpose is to provide continuous feedback to both student and teacher concerning learning successes and failures. Feedback to students provides reinforcement of successful learning and identifies the specific learning areas and misconceptions that need correction. Feedback to the teacher provides information for modifying instruction and for prescribing group and individual work.... Because formative assessment is directed toward improving learning and instruction, the results typically are not used for assigning course grades. (Linn & Gronlund, 2000, p. 41)

In order to deliver the greatest benefit, formative assessment should begin as soon as possible after the commencement of a course and should be an ongoing process. Regular and ongoing, formative assessment can target in depth small parts of the course in a way that summative assessment cannot:

> if evaluation is to aid both the teaching and learning processes, it must take place not only at the termination of these processes but while they are still fluid and susceptible to modification. Formative evaluation, as the name implies, intervenes during the formation of the student, not when the process is thought to be completed. It points to areas of needed remediation so that immediately subsequent instruction in study can be made more pertinent and beneficial. Formative evaluation impinges on smaller, comparatively independent units of the curriculum... assessing in depth a relatively small number of objectives. (Bloom et al., 1971, p. 20)

As well as offering guidance to students and teachers, formative assessment brings further benefits. Properly managed, it can provide encouragement to students. In addition, it can offer students the opportunity to become familiar with the sort of assessment that will be used summatively later in the course (Charman, 1999).

In spite of the clear benefits of formative assessment, the concept appears to have been isolated and named only relatively recently in the history of education. It seems that the term 'formative' was first used by Scriven in 1967 but in the context of curriculum evaluation. Scriven made the point that once a curriculum has been put in its final form, it is extremely difficult to change. A better curriculum would result if evidence of the appropriateness of the curriculum could be collected and acted on during its construction. Bloom, Hastings and Madaus applied this notion and terminology to the area of student learning in 1971. While the term is now used widely in assessment literature, it has been noted that it does not have a strictly defined meaning (Sly & Rennie, 1999). In fact, one recent text appears to strongly relate "formative" to "informal", although the description given does cover approximately the same ground as descriptions from other commentators:

> Formative, or informal, assessment occurs as you proceed with a lesson or unit to provide ongoing feedback to the teacher and student.... The purpose of formative assessment is to provide corrective actions as instruction occurs to enhance student learning. It consists of informal observation, questioning, student responses to questions, homework, worksheets, and teacher feedback to the student. Formative assessment is integrated with instruction on a daily basis. It is flexible and idiosyncratic. By contrast, summative, or formal, assessment takes place at the end of a unit of study. (McMillan, 2004, p. 105)

While much formative assessment is informal in nature, it need not always be so, particularly in the larger classes found in universities. The key feature of formative assessment is that it occurs at the formative stage, helping the learner better form her or his skills, knowledge, learning patterns, and so on, not that it is informal.

The research literature contains many examples of studies in which implementation of formative assessment led to improved student achievement (Buchanan, 2000; Charman & Elmes, 1998). At the University of Luton, Zakrzewski and Bull (1998) found that the adoption of formative assessment in a biology course led to an average increase of 1.6 grade points in the final examinations. It was also noted that examination stress was greatly reduced: "Any student anxiety had been overcome by previously using the system in a formative way" (p 149).

Gretes and Green (2000) studied over 200 students in an undergraduate education course. Practice tests were offered as an option, with two-thirds of the students sitting one or more practice tests (average: three tests). The practice tests assessed the same course objectives that were later assessed by summative assessment, but contained entirely different items. The students who did opt for practice tests scored more highly in the final examination: 81% versus 76%. This difference was significant at $p<0.001$. An attitude survey also indicated that the students viewed the practice tests favourably. Ninety three percent of the students agreed or strongly agreed with the statement: "the practice test helped me study for the real test", while 81% of students reported that they gained "more confidence" for taking the graded in-class exam. A second study was carried out to support and extend the results of the first study. In this, 566 students, aged from 19 to 52 were studied. Seventy five percent of the students were females. Again, an opportunity to take part in computerised practice testing was given one week before the exams. Sixty percent of the students took advantage of this, averaging 3.89 test-taking events. Those who did take the computerised practice tests had an average exam result of 84.2% versus 77.2% for those who did not. It was also found that there was a significant relationship between the number or practice tests taken and the average exam result even after the covariant of SAT score was controlled ($p<0.005$).

The division of Electronic and Electrical Engineering at the University of Northumbria offered their students the option of practice tests. The take-up of the voluntary self-tests was very high and appeared to shift learning patterns from surface learning to something deeper: "We found evidence that the practice tests encouraged students to study effectively, putting effort into understanding the course content as the unit progressed, rather than 'cramming' for exams at the last minute" (Sambell et al., 1999, p. 183). The implementation of practice tests also led some students to view tests in a different light, not as an obstacle but, rather, as a learning aid: "Some saw the practice tests as 'friendly' because they interpreted that the tests were there to help them learn more effectively rather than test them" (p. 184). "Most valued the practice tests because they helped them, personally, to identify the areas for improvement, thus helping them direct future effort in fruitful directions" (p. 185). Others in the area of Engineering have also commented on the way formative assessment can assist students to develop deeper learning practices: "To encourage deep learning the emphasis should be put on formative evaluation, that is, the posing of problems to students and the provision of adequate feedback to encourage reflection, discussion and critique" (Fernandez et al., 2001, p. 333).

A study carried out at Curtin University of Technology involving several hundred students found that students taking a practice test outperformed students who did not even though, on the basis of their entry scores, there should have been no significant difference. The difference was statistically significant at the $p<0.001$ level and persisted through similar follow-up studies, even when apparently weaker students opted to take the practice test (Sly, 1999). Two explanations for improvement following the practice test are offered:

> first, that students use the feedback they receive on incorrect responses to assist in their learning of the topic and, second, that increased familiarity with the CML [Computer-managed learning] system may contribute to a better performance the next time. (Sly & Rennie, 1999, p. 119)

The researchers also felt that feedback from the practice test could have been of valuable assistance to instructors, not just to students. However, they saw little evidence of it being used.

At Michigan State University (MSU) a novel approach was employed that saw mid-term examinations being used in a formative way. After sitting the examination, students were offered the option of re-sitting a personalised and different version of the original examination. This re-sitting was carried out online under the same conditions that were used for homework assignments. Thus, students had access to hints, could participate in discussion forums, and could obtain help from instructors. It was found that virtually all students made use of the opportunity. The incentive was the potential to increase one's marks: 25% of marks previously missed could be recovered by solving all the problems correctly. Thus a student who scored 60% on a mid-term exam and gained 100% on the version following the exam would subsequently earn a grade of 70%. It was noted that the system would have been unmanageable without the use of computers. The software used was CAPA: Computer-Assisted Personalized Approach. One aspect of the formative nature of the re-sit experience can be seen from the way student attitudes to the examination reportedly changed on taking the re-sit:

> An interesting impact of this 'Partial Credit by Corrections' is that the number of students complaining about the difficulty and/or unfairness of the exam drops considerably after they get the correct solutions. A much more common response is 'It wasn't that hard! I should have got it!' (Albertelli, Kortemeyer, Sakharuk, & Kashy, 2003) (p. 6)

A landmark review of the literature on classroom formative assessment was carried out by Black and Wiliam (1998a, 1998b). The authors looked over a large number of substantial review articles, carried out key-word searches in the ERIC database, scanned 76 of the most likely journals for the last decade, and followed up references in articles already found. They typically found positive effect sizes of between 0.4 and 0.7 in formative assessment experiments (Black & Wiliam, 1998b). It was noted that an effect size of 0.7, if it could be achieved on a nationwide scale, would be equivalent to raising the mathematics attainment score of an 'average' country like England, New Zealand or the United States into the 'top five' after the Pacific rim countries of Singapore, Korea, Japan and Hong Kong: "significant learning gains lie within our grasp. The research reported here shows conclusively that formative assessment does improve learning. The gains in achievement appear to be quite considerable… amongst the largest ever reported for educational interventions" (Black & William, 1998a, p. 61).

These learning gains made through formative assessment could be undermined by pre-occupation with activities associated with summative assessment such as giving grades, and other normative activities. Task-involving evaluation was seen as more effective than ego-involving evaluation. It was also noted that "the learning goal orientation led to higher motivation and achievement outcomes than did the performance goal" (Black & Wiliam, 1998a, p. 14).

In an earlier review of the literature, Crooks (1988) also concluded that formative assessment could be a most powerful force:

> classroom evaluation affects students in many different ways. For instance, it guides their judgment of what is important to learn, effects their motivation and self-perceptions of competence, structures their approaches to and timing of personal study (e.g., spaced practice), consolidates learning, and affects the development of enduring learning strategies and skills. It appears to be one of the most potent forces influencing education. Accordingly, it deserves very careful planning and considerable investment of time from educators. (p. 467)

However, he concluded that such planning and investment of time was not taking place. Rather, the emphasis was on summative assessment: "too much emphasis has been placed on the grading function of evaluation, and too little on its role in assisting students to learn" (Crooks, 1988).

Of course, formative assessment can only lead to significant learning gains when the learner uses the information it provides to his or her advantage. This requires two actions on the part of the learner. The first is the perception of a gap between a desired goal and her or his present state. The second is the action taken to close that gap in order to attain the desired goal. For the first action, students need to be aware of the desired goal, and they also need to be able to compare their work with it:

> for students to be able to improve, they must develop the capacity to monitor the quality of their own work during actual production. This in turn requires that students possess an appreciation of what high-quality work is [and] that they have the evaluative skill necessary for them to compare with some objectivity the quality of what they are producing in relation to the ... standard. (Sadler, 1989, p. 119)

Clearly, the student has more than a passive role in this, particularly if the formative assessment is in the form of an informal or optional self-assessment. Even when the formative assessment is a compulsory assessment, the student will need to draw

some conclusions from the result concerning the gap between his or her present and desired standards. This involves much more than simply comparing an actual score with a desired score: for example, decisions will also need to be made concerning the nature of the learning activities pursued and the amount of effort devoted to them. Following these decisions, the student must then go on to choose and implement a course of action appropriate to reducing the gap between present attainment and future goals. Self-assessment on the part of the student is seen as an essential ingredient for learning (Black & Wiliam, 1998a; Sadler, 1989). The role of the instructor in formative assessment is also critical. It goes beyond setting up the assessment and providing the student with information concerning current and desired standards, even though these in themselves are daunting tasks. The instructor is also charged with the delicate task of setting up a learning environment such that, when the students receive information concerning the gap between actual and desired levels of achievement, they have the knowledge and ability to choose appropriate courses of action and the motivation to implement them effectively. One aspect of motivating students will be inculcating in them "the idea that success is due to internal, unstable, specific factors such as effort, rather than on stable general factors such as ability (internal) or whether one is positively regarded by the teacher (external)." (Black & Wiliam, 1998a, p. 51) These many factors involved in successful formative assessment make it difficult to determine why formative assessment does not lead to improved summative assessment results in those rare situations where it does not (Peat & Franklin, 2003).

2.5 The affects of assessment on students

As more research is conducted, the pivotal role played by assessment in the learning process becomes even clearer. For example, research at Oxford Brookes University in the mid-1990s suggested that differences in student performance are tied, not to class size as was previously thought, but to assessment regimes (Gibbs & Lucas, 1997). Certainly, the considerable evidence for the motivating power of assessment has been well-established for some time: "The quickest way to change student learning is to change the assessment system" (Elton & Laurillard, 1979, p. 100), "student motivation to engage in activities not contributing directly to final marks is low" (Rothberg, Lamb, & Wallace, 2001, p. 507), "The principle that 'what you test

is what you get' is widely appreciated in education" (Raikes & Harding, 2003, p. 267). Extensive study of student diaries at Leeds Metropolitan University (Innis, 1996) revealed that, by year four, only 5% of out-of-class student time was spent on learning unrelated to assessment. Gibbs (1999) suggests that assessment is the most powerful lever teachers have to influence the way students behave as learners.

Snyder (1971) in 'The Hidden Curriculum' commented that the assessment system implemented in a course heavily influences the way students approach their learning. It affects the amount of effort they make and where they concentrate their efforts: it acts as a 'hidden curriculum'. Others have made similar comments: "examinations tell them our real aims, at least so they believe" (Rogers, 1969, p. 956), "It is well known that assessment arrangements define the curriculum in the eyes of the learner" (Mason, 1998, p. 41). Assessment subliminally sends strong messages to students not only about the course aims and content but also about the type of learning regarded as important (Sambell & McDowell, 1998; Wolf, 1993).

Crooks (1988) lists ways in which assessments affect students. He classifies these effects into three groups based on time perspective. For the shortest interval, such as a particular lesson, topic, or assignment he includes:

> consolidating prerequisite skills or knowledge... focusing attention... encouraging active learning strategies... giving students opportunities to practise skills... providing knowledge of results and corrective feedback... helping students to monitor their own progress and develop skills of self-evaluation... guiding the choice of further instructional or learning activities to increase mastery... helping students feel a sense of accomplishment. (p. 443)

For the medium term, such as a particular learning module, course or extended learning experience, Crooks sees the following as important effects of evaluation:

> checking that students have adequate prerequisite skills and knowledge... influencing students' motivation to study the subject... communicating and reinforcing broad goals for students... influencing students' choice of (and development of) learning strategies and study patterns... describing or certifying students' achievements in the course. (p. 443)

Finally, longer term effects of assessment include:

> influencing students' ability to retain and apply in varied contexts and ways the materials learned... influencing the development of students' learning skills and styles... influencing students' continuing motivation, both in particular subjects and more generally... influencing the students' self-perceptions. (pp. 443 - 444)

There is a great deal in the above. For example, helping students "monitor their own progress and develop skills of self-evaluation" (p. 443) and "influencing students' motivation to study the subject" (p. 443) resonate with the very important goal of bringing students to the point where they can take charge of their own learning so that their learning does not effectively cease once they leave school or university. If assessment helps students in this way, and it is argued that it can (Bell & Cowie, 2001), then much will have been achieved.

Focussing on the important area of end-of-topic tests, Crooks (1988) found numerous studies showing that such a test promoted longer-term retention of the material studied, even when no feedback was given on the test performance. Crooks gave three reasons for this. First, the test motivated the students to familiarise themselves with the material another time, a limited form of distributed practice. Second the test encouraged the students to actively process content. Third, the test directed the attention of the students to the topics, skills, and details tested.

Assessment can also be used to motivate students to attend lectures. For example, at MSU individualised, pre-printed quizzes are completed by students during the last part of a lecture (Albertelli, Kortemeyer, Sakharuk, & Kashy, 2003). These quizzes, because they make a small contribution to the final grade, motivate the students to attend lectures and pay close attention to the lecture. The quizzes are machine-generated and marked. Without machine assistance, such an assessment programme would not be viable.

The motivating power of assessment persists even when students rather than the instructor carries out the assessment. In fact, in some situations considerable gains can be made by replacing instructor assessment with peer assessment. Gibbs (1999) describes a compulsory second-year engineering course where the format was two lectures and one tutorial a week. At the tutorial students worked on problem sheets which were marked by lecturers and handed back each week. The average exam mark was 55%. Numbers increased and the lecturers no longer had time to mark the problems. The average exam mark went down to 45%. In an effort to restore standards, peer assessment was implemented. Six times during the course, students

met and handed in their problem sheets. These were then redistributed randomly with a mark scheme. Students then marked the work they were given and the papers were returned immediately. The marks were not recorded and teachers were not involved at all in this process. Students were required to complete about three quarters of the problem sheets, otherwise they could not complete the course. Everything else about the course remained the same. Amazingly, the average exam mark went up to 75%. There are two conclusions that may be drawn from the above. First, the weekly problem sheets, when marked, motivated students to actually solve problems and distribute their effort more effectively rather than concentrate it in the period just prior to the final examination. It is also likely that having one's work marked by peers is more daunting than having it marked by some relatively unknown tutor. However, it is improbable that this in itself would have produced such a large achievement gain. Second, there is a clear benefit from having students actually carry out the marking process. The act of examining a mark schedule and following it gives a fuller perspective than simply doing the problems can give. In marking the problems of other students, the markers would have noticed other ways to do the problems correctly, errors which they made themselves, and errors which they were alerted to avoid. They would also see the care some students took and others did not. They would get a good idea of the standard of work and the effort required to achieve at various levels. In time, the process of marking enables students to internalise the appropriate standards. Once the standards are internalised, students are equipped to be able to monitor the quality of their own work. The importance of this is illustrated by the large lift in achievement that occurred when peer marking was introduced. However, students do not always appreciate the benefit to themselves to be gained by marking the work of their peers: “When students object to self-assessment or peer assessment it is often because they do not understand the importance of internalizing standards in order to be able to supervise one's own work” (Gibbs, 1999, p. 47). Mafi (1989) also reported increased student understanding and involvement, together with decreased instructor workload, as a result of implementing peer assessment in civil engineering.

Of course, not all the effects of assessment are positive. Since assessment has such power over the way in which students approach their studies, it can encourage both desirable and undesirable behaviours. For example, Crooks (1988) comments that

students' approaches to learning tasks can be broadly categorised as either deep or surface approaches, and that they appear to be somewhat versatile in their choice of learning approach. While the choice depends on a number of factors, such as how interesting the topic is and how much time and energy are available, the dominant factor appears to be the method of evaluation. Students who display deep approaches to learning have the intention to understand what they study. This involves being able to organise and structure knowledge, to see relationships between different areas, and to retain and use the knowledge being acquired. By contrast, students adopting a surface approach to learning are driven by the goal of passing the assessment system. They are extremely task-focused, learn unreflectively, and rely largely on memorization and rote learning.

> Students who are judged to be too extrinsically motivated are often labelled as 'assessment driven' and the behaviour they exhibit can be seen as a problem. This is a commonly observed phenomenon in medical education, where students are observed to only learn those elements of the curriculum that are known to be directly assessed and to be more concerned about grades than about using assessments as a learning experience. (Seale, Chapman, & Davey, 2000, p. 615)

Assessment methods can predispose students to choose either learning approach. Surface approaches are encouraged by assessments that cover excessive amounts of material, emphasize recall and provoke anxiety. Unfortunately, as instructors struggle to keep their courses up to date (an ongoing problem in engineering and many other areas) they are likely to add in course material without pruning out an equal amount of older material. In this way courses tend to become larger rather than smaller: faced with excessive course material, students may be encouraged into surface learning approaches (Berry, DiPiazza, & Sauer, 2003; Soderstand, 1994). Deep approaches are encouraged by assessment methods which require active engagement with learning tasks: understanding concepts and making connections (Bell & Cowie, 2001). When the assessment of a course is concentrated in one final examination, students may be tempted to put off work till the end of the course and then cram at the last minute. Such an approach is unlikely to promote understanding. The sorts of questions used in the examination also affect how students prepare for it and, more generally, how they approach the course: questions that are largely mathematical manipulations will engender a very different approach to questions that

are largely descriptive (McDowell, 1995). Black and Wiliam (1998a) see the problem of assessment practices encouraging surface learning as widespread:

> classroom evaluation practices generally encourage superficial and rote learning, concentrating on recall of isolated details, usually items of knowledge which pupils soon forget… Teachers do not generally review the assessment questions that they use and do not discuss them critically with peers, so there is little reflection on what is being assessed... The grading function is over-emphasised and the learning function under-emphasised. (p. 17)

Crooks (1988) notes that norm-referenced evaluation produces a number of undesirable consequences such as

> reduction of intrinsic motivation, debilitating evaluation anxiety, ability attributions for success and failure that undermine student effort, lowered self-efficacy for learning in the weaker students, reduced use and effectiveness of feedback to improve learning, and poorer social relationships among the students. (p. 468)

Assessment, particularly normative assessment, can also produce a competitive environment. In general, such an environment produces lower achievement than a cooperative one. In particular, the weaker pupils tend to get the message that they lack ability, so that they are de-motivated and lose confidence in their own capacity to learn (Black & Wiliam, 1998a). Crooks (1988) also notes that when students who are initially intrinsically motivated are confronted with attempts to stimulate their learning through extrinsic motivation, the usual result is decreased intrinsic motivation, especially on challenging tasks. He concludes:

> for many students schooling tends to lower rather than increase interest in learning ... the key factor seems to be whether students perceive the primary goal of the evaluation to be controlling the behaviour or providing information and helpful feedback on their progress in learning. (p. 464)

In the light of the above, summative assessment clearly can be problematical. On the other hand, formative assessment, because it provides feedback on student learning, has the potential to increase student interest in learning: "To encourage deep learning the emphasis should be put on formative evaluation, that is, the posing of problems to students and the provision of adequate feedback to encourage reflection, discussion and critique" (Fernandez et al., 2001, p. 333).

Crooks (1988) found that a large number of studies showed substantial negative correlations between test anxiety and test performance. The correlation was stronger at higher grade levels. Test anxiety was likely to be particularly high "when the

student perceives good performance on the test to be particularly important, when the test is expected to be difficult, and when the testing conditions are particularly intrusive" (Crooks, 1988, p. 461). In order to reduce test anxiety, and therefore promote test performance, instructors should make a practice of having very generous time limits, giving the students ample information about the test, setting tasks that are not unreasonably difficult, and de-emphasising social comparison (Crooks, 1988). One way students can gain more information about the test difficulty, nature and format is to sit a practice test. This is frequently done in the case of computer testing: the increased familiarity the students gain with the test environment reduces test anxiety to the extent that it no longer significantly affects their performance (Sly & Rennie, 1999; Zakrzewski & Bull, 1998).

2.6 Frequency of assessment

There are arguments both for and against increased frequency of assessment. In support of more frequent testing, one may argue that more tests would encourage students to study and review more often, provide more regular feedback, and give students a good indication of what they were expected to learn. Against more frequent testing, it may be argued that more tests would leave less time for instruction. Further, students might focus on passing tests rather than on learning. Too frequent testing could become monotonous and reduce enthusiasm for learning. Because tests tend to act as signals to students that one section of work has finished and another is about to begin, courses that should be understood as an integrated whole might become fragmented in the minds of students by too-frequent testing.

Bangert-Drowns, Kulik and Kulik (1991) reviewed evidence on the effects of frequent testing. They carried out a meta-analysis of 40 relevant studies and found that performance did improve with frequent testing. However, it was a case of diminishing returns: "taking one test during a 15 week term seemed to provide almost as much preparation for a criterion examination as did higher test frequencies" (p. 94). The actual figure from their regression analysis was 0.34 standard deviations. Increasing the number of tests per term from one to two increased student achievement by only 0.07 standard deviations, and further increasing the number of tests per term from two to four increased student

achievement by only 0.08 standard deviations. However, it should be noted that all the research surveyed in the meta-analysis was carried out in classrooms. In the classroom situation there is a teacher present to keep students on-task. This does not happen in the university situation. Therefore, it is possible that more frequent testing produces better results at university level than it does at lower educational levels.

Black and Wiliam (1998b) state that "it is better to have frequent short tests than infrequent long ones. Any new learning should first be tested within about a week of a first encounter, but more frequent tests are counterproductive" (p. 144). On the basis of their analysis mentioned above, Bangert-Drowns et al (1991) would certainly agree that testing more than once a week would be counterproductive. However, they would most likely suggest that the testing of material could take place more than a week after the first encounter with that material without a significant drop in student learning. The argument that frequent short tests are better than infrequent long ones also needs closer scrutiny. In another meta-analysis, Bangert-Drowns et al (1991), considered studies where the same items were used in differentially-distributed tests during instruction. In some classes the items were distributed in short, frequent quizzes, while in other classes the items were distributed in longer, less frequent quizzes. It was found that "The students performed better on short quizzes on short units of instruction than on longer quizzes over the longer units of instruction" (p. 96). The difference was 0.57 standard deviations. However, this difference did not translate into better post-test performance: "Nine studies with differential distribution of identical items reported the performance of students on post-tests... The average effect size comparing the more frequently tested group to the less frequently tested group was zero" (pp. 96-97). The inference is that the benefits of more-frequent testing are short-lived. Perhaps it is simply an illustration of the fact that it is easier to study for a test when there is less to study for it. A meta-analysis of four studies of student attitudes to testing concluded that students preferred classes in which the testing was more frequent. A large average effect size of 0.59 was noted (Bangert-Drowns, Kulik, & Kulik, 1991; Crooks, 1988). None of the studies was recent: the most recent was 1981, while the oldest was 1968. The most recent of these looked at student preferences regarding test frequency, how test frequency affected student performance, and test anxiety. 400 psychology students at Western Illinois

University were involved. 200 students had an objective 25-item test every two weeks, while the other 200 had an objective 50-item test every four weeks. The same test items were used for each group. The research findings were in agreement with studies discussed previously: a higher frequency of tests produced better test results but no final examination advantage, students preferred more frequent tests, and highly anxious students did less well than low anxious students:

> frequent testing helps students to perform better on a test-by-test basis, but this advantage disappears on an announced final exam when, presumably, cramming can equalise knowledge of course material. Students definitely favour frequent tests over infrequent tests as indicated by both subjective ratings of the instructor as well as by subjective comments. In addition, highly anxious subjects do better on short frequent tests than on longer infrequent ones, although they still perform below the level of equivalent low anxious subjects on either kind of test. Finally, the relative improvement of high anxious subjects on short frequent tests does not carry over to the performance on an announced final exam. (Fulkerson & Martin, 1981, p. 93)

2.7 Assessment format

Increased staff workloads and the upsurge of the use of computers in education are driving a move towards objective assessment items that can be marked either automatically by a computer or by an assistant who may not be able to bring to the marking process the fine judgment that the course instructor possesses (Sly & Rennie, 1999). For example, at the University of Luton, there has been a significant shift to final examinations that are delivered by computer and are largely multiple-choice: a typical final exam is one hour long and contains 60 to 80 questions (Zakrzewski & Bull, 1998). The software used is 'Question Mark Designer' for Windows (Questionmark, 2004).

Excell (2000) studied the results of introducing multiple-choice examinations for topics such as electromagnetics and optics. No major anomalies were noted in the marks scored by students:

> it was found that the mark distributions were comparable with those achieved by the same groups of students in traditional examinations and there appeared to be a good correlation with the abilities of the students, as measured in the traditional way. Fears that the examination would be too superficial and that students would easily score high marks proved groundless. (p. 250)

However, Excell did accept that multiple-choice questions might not explore an issue as deeply as some other sorts of questions. This possible disadvantage he saw as being balanced by advantages such as being able to cover the whole of the syllabus. Another disadvantage is the increased possibility of cheating and the relative difficulty of detecting such cheating. To counter this, a number of institutions are now scrambling or randomising their examinations (Harpp, 1994). This is particularly easy to do when CBA is used (Thoennessen, Kashy, Tsai, & Davis, 1997).

Students sometimes comment negatively about the element of luck involved in answering multiple choice questions (O'Reilly, Alexander, Sweeney, & McAllister, 2002). One perceived difficulty is the way in which marks can be scored merely by random guesses. A response to this is to implement a system designed to make guessing unrewarding. For example, for multiple-choice questions having five options, each correct answer could be awarded +1 while each wrong answer receives -0.25. However, it has been demonstrated (Noll, Scannell, & Craig, 1979) that such a system has no significant effect on the rank-ordering of the students. Excell (2000), being aware of the above but wanting to not reward students who merely guessed answers, implemented the 'stupid answer' method at Bradford University. In the simplest form of this method, the right answer scores +1. One answer, regarded as being obviously wrong to anyone who has at least attended the lectures, scores -1. The other plausible but incorrect answers score 0. The method appeared to work well when tried with one electromagnetics course and the suggestion was made that its use would be extended into other courses. In the future, multiple-choice testing may become more prevalent at university level and more sophisticated marking schemes such as this could play an important role in assessment.

Crooks (1988) takes a neutral stance on the issue of test format. He sees no strong evidence "to support widespread adoption of any one item format or style of task. Instead, the basis for selecting item formats should be their suitability for testing the skills and content that are to be evaluated" (p. 448). Paxton (2000) believes that multiple-choice testing is all too prevalent, particularly in some large first-year classes, and argues for the use of multiple-choice questions as part only of a broader and more diverse range of assessment measures. This would enable students to

develop communicative competence in the discourses of the academic disciplines they are studying. Fowell, Southgate, and Bligh (1999) also favour a diverse range of assessment methods since particular methods may disadvantage particular students. They suggest that using a selection of methods allows better triangulation in determining student achievement. Seale et al. (2000) investigated which types of assessment students found most motivating for their learning. They concluded that the students responded positively to a wide range of assessments. Students also appreciated the variety of assessments because it gave them an opportunity to express many different facets of performance. Against this, some students felt that the variety did not give them adequate opportunity to build on their skills. It was also found that the relevance and content of the assessment appeared to influence student motivation as well as the enthusiasm of the teachers.

2.8 Summary

Assessment may be summative or formative. Summative assessment primarily aims to grade students. Formative assessment aims to provide feedback to students and teachers about learning progress. Traditionally, summative assessment has been overemphasised. The research literature overwhelmingly argues that more emphasis on formative assessment enhances learning: lifting examination achievement and shifting learning patterns from surface learning and last-minute cramming to something deeper as students place more emphasis on understanding. This is best promoted through frequent, short assessments that start early in the course.

OASIS provides instructors with the ability to deliver such assessments. Typically, the first OASIS assignment takes place in the second week of a course and assessments are fortnightly thereafter. In line with the wisdom that a diverse range of assessment styles is best, traditional written tests are also included in the mix. Instructors who have introduced OASIS into courses do report enhanced student achievement as measured by examination results and by student feedback. In line with good formative assessment practice, OASIS motivates students to study without provoking the extreme anxiety associated with high-stakes summative assessment; rather, students frequently see it as a learning aid.

Chapter 3. Review of the Related Literature: Feedback

3.1 Purpose and outline of chapter

Whereas Chapter 2 examined the literature relating to assessment in general and Chapter 4 will consider the literature on CBA, the present chapter reviews the literature relating specifically to feedback. The nature of feedback systems is outlined in Section 3.2 and the effect of such systems on students is examined in Section 3.3. The next three Sections focus on specific issues: the argument against pre-search availability (Section 3.4), the notion that extensive feedback may not be more effective than feedback that supplies only the correct answer (Section 3.5), and the importance of prompt feedback (Section 3.6). Finally, a summary of the chapter is presented in Section 3.7.

3.2 Introduction

The purpose of formative assessment is to provide feedback to instructors and their students in order to support future learning. Therefore feedback is the very essence of formative assessment: "In order to learn, students need to know what they don't know!" (E. Kashy, Thoennessen, Tsai, Davis, & Wolfe, 1997), "Knowledge of results plays a major role in almost all learning and motivational theories of task performance" (Ilgen & Davis, 2000, p. 550),

> Any theory that depicts learning as a process of mutual influence between learners and their environments must involve feedback implicitly or explicitly because, without feedback, mutual influence is by definition impossible. Hence, the feedback construct appears often as an essential element of theories of learning and instruction. (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991, p. 214)

Feedback does not play such an important part in summative assessment. For example, the information that one scored a grade of C in a final examination does not offer much guidance for future study. In this respect it is on a par with the exhortation: "try harder!" More feedback could be provided following summative assessments. It would be possible, for example, to report to the student how well they performed in various aspects of the assessment. At UoA, students can request a copy of their final examination script and from this can gain some information about how they performed on the various questions. However, markers write no comments on

the script, simply the marks gained for each question. Even extensive feedback following summative assessment may be of little benefit to the student concerned: if the student is not sitting the paper, or one similar to it, again there may be little motivation to learn from the feedback. In fact, some commentators (Sadler, 1983; Slavin, 1978) suggest that the feedback and summative purposes of evaluation are best separated. Crooks suggests that "where evaluations count significantly toward the student's final grade, the student tends to pay less attention to the feedback, and thus to learn less from it" (p. 457). This literature review will therefore naturally focus on feedback in the context of formative assessment.

Black and Wiliam (1998a) identify four elements that should make up every effective feedback system:

- data on the actual level of some measurable attribute;
- data on the reference level of that attribute;
- a mechanism for comparing the two levels, and generating information about the gap between the two levels;
- a mechanism by which the information can be used to alter the gap (p. 48).

Ramaprasad's (1983) definition of feedback captures these same four elements: "Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way" (p. 4). This definition of feedback contains precisely the same key elements that make up the concept of formative assessment. Thus, the research surveyed previously that found heavily in favour of formative assessment also provides a strong argument for feedback.

However, not all feedback would seem to fit the above definition. A student who attempts an on-line question and receives a score of 2 out of 5 may not be sure whether she or he has attained the reference level or not. Even scoring full marks for a very easy question may not, in itself, reassure a student that he or she has reached the required standard. A remark such as "good work" made by a tutor during a tutorial also lacks some of the four elements above. In the light of these examples, perhaps a more broad view of feedback should be employed. Kluger and DeNisi (1996) define feedback interventions as "actions taken by (an) external agent (s) to provide information regarding some aspect (s) of one's task performance" (p. 255).

This definition contains only the first of the four elements above. It also is broader than the concept of 'knowledge of results'. For example, "you do not use your thumb for typing" is not a 'knowledge of results' intervention but it is a feedback intervention. However, this definition does exclude situations where the feedback is self-generated. Kluger and DeNisi state, for example, that a gardener noting that she or he has flooded a plant is not an example of a feedback intervention.

3.3 The effects of feedback on students

Lack of feedback on progress has been advanced as an important contributing cause of failure of first year students (Entwistle, Hounsell, Macaulay, Situnayake, & Tait, 1989). McDowell (1995) suggests that lack of feedback to students can lead to "ineffective learning, poor use of time, anxiety and lack of motivation... this issue is even more serious for students who have failed in the past and may lack confidence" (p. 419).

Kluger and DeNisi (1996) reviewed over 3000 reports on the effects of feedback interventions. They discounted most of the reports on the basis of lack of rigour or lack of detail. A meta-analysis of the remaining 131 reports (607 effect sizes and 23,663 observations) yielded an average effect size of 0.41 (equivalent to raising the average student to the 65th percentile). However, the standard deviation in the effect sizes was almost 1, and over one third of the feedback interventions actually decreased performance. They found little acknowledgement in the literature of just how frequently feedback interventions produce decreased performance. Somehow feedback is seen as desirable. Is this because it is human to want feedback? Or is it because feedback interventions have other benefits? For example, feedback may increase satisfaction and it may contribute to long-term persistence on the task.

Kluger and DeNisi argue that it is too simplistic to explain behaviour in terms of the gap between actual and reference levels, and discrepancy reduction. In order to explain how feedback interventions (FIs) can produce negative effects on performance, they suggest that

> FIs change the locus of attention among 3 general and hierarchically organized levels of control: task learning, task motivation, and meta-tasks (including self-

> related) processes. The results suggest that FI effectiveness decreases as attention moves up the hierarchy closer to the self and away from the task. (p. 254)

Black and Wiliam (1998a) and Ames (1992) similarly make the point that task-involving feedback is more effective than ego-involving feedback. The awarding of grades has ego-involving aspects and can produce negative motivational effects. Crooks (1988) also suggests that there is little research support for making praise a prominent part of feedback although he regards simple positive comments as beneficial and harsh criticism as counter-productive. He notes that younger and less able students may benefit most from praise. Against this, however, Black and Wiliam (1998a) state that praising low-achievers can have a negative effect. Feedback should give each pupil specific guidance on strengths and weaknesses. It should not include any overall marks, nor any comparisons with other students (Black & Wiliam, 1998b). Crooks (1988) concurs: "feedback is most effective if it focuses students' attention on their progress in mastering educational tasks. Such emphasis on personal progress enhances self-efficacy, encourages effort attributions, and reduces attention to social comparison" (p. 468). Ilgen and Davis (2000) believe that the most critical issue in effectively delivering negative feedback is achieving student responsibility for substandard performance without lowered student self-concept. They see focusing on learning rather than on performance as a way to accomplish this. Minimising notions of interpersonal competition is also seen as likely to be helpful.

Feedback fails to perform its facilitative role in two different types of situations. One of these occurs when the material studied is very difficult for the learners. In this case they are likely to simply spend most of their time guessing the answers, and then trying to see how the supplied answer corresponds with the question (Kulhavy, 1977). The second situation concerns 'pre-search availability'.

3.4 Pre-search availability

'Pre-search availability' refers to the situation in which students can see the answers before they have attempted the questions. While it is sensible to make available to students the correct answers for the questions that they are working on, this raises the possibility that they can look at the answers before seriously tackling the question;

when the questions are from a text that has the answers at the back, for example. Where pre-search availability exists, the feedback associated with providing the correct answers loses its effectiveness (Black & Wiliam, 1998a). Kulhavy (1977) similarly noted that, with pre-search availability, feedback often had an insignificant or even detrimental effect on achievement. By contrast, when pre-search availability was controlled, students that received feedback demonstrated higher achievement than peers who received no feedback. When it was not controlled, however, students tended to copy answers and avoid studying the material at all. The copying was indicated by the fact that the students made fewer errors during instruction and completed tasks more quickly yet gained lower achievement scores in final testing.

Bangert-Drowns, Kulik, Kulik et al., (1991) came to the same conclusion. Pre-search availability was seen as encouraging mindless copying of correct answers. It led to far fewer errors being made during practice, but this did not carry through to the final assessment. Providing feedback with pre-search availability led to slightly lower achievement than providing no feedback at all (average effect size = -0.08). When feedback without pre-search availability was provided, it made an important contribution to achievement (average effect size = 0.46). Clearly, control for pre-search availability was strongly related to effect size ($F[1,56] = 22.77$, $p = 0.0001$).

Some instructors have noted that pre-search availability can produce students who become too reliant on the feedback and use it as a substitute for their own efforts: "initial experience with the feedback system showed that students depended too heavily on the feedback system as a substitute for their own testing" (Chen, 2004), "the fact that the solution was readily available if needed meant that they did no work at all... They never attempted to complete any of the tasks and relied on the system to present them with the solution" (Davidovic, Warren, & Trichina, 2003).

It is not a simple task to decide to what extent students will access answers in advance of attempting questions nor is it straightforward to determine the subsequent effect on their learning if they do. For example, consider the situation where students are set some questions from a text and the answers are in the back of the text. It is likely that some students will access the answers before they have made any significant progress on the questions, some students will work on the questions for a

while and then look at the answers for guidance if they feel they are not making much progress, and some students will steadfastly refuse to look at the answers until they have produced their own. It is extremely difficult to determine in advance what courses of action students will choose. The choices made will depend on the motivation of the students at the time, how busy they are, how challenging the problems are, and the worth of that challenge. These factors are not absolutes either, but vary from student to student and are to a significant extent a product of the perceptions of the students. Another factor is the ease of gaining access to the answers. A software package that allowed the answers to be viewed by simply clicking on a button might be just too tempting for most students. On the other hand, if students had to submit their own answers before gaining access to the correct answers then fewer would be tempted. Another factor to consider is the perceived value of the answer itself. For example, a question which asks students to name the capital of some country would be very susceptible to pre-search availability. This is because the question really entails nothing more than knowledge of the answer. By contrast, a question requiring the solution of some relatively complex electrical circuit may be much less susceptible to pre-search availability. This is because the question is likely to be perceived as testing one's ability to choose and correctly apply a method to solve the circuit, rather than one's ability to magically produce the correct answer. The correct answer is only provided as a quick check that one has properly applied an appropriate method to solve the circuit. Looking at such an answer in advance may give no clues at all about how to actually solve the circuit. Finally, there is the issue of how much importance is actually attached to the correct answers. If the students have to submit to the instructor the correct answers then they may well copy the answers rather than derive them themselves. If, on the other hand, the problems have been set merely for the purpose of practice and the students are not required to submit their answers, then they are much less likely to resort to copying. Of course, it can also be argued that, in this situation, some students will be less likely to attempt them at all!

3.5 Extent of feedback

Three main possibilities are considered here:

(1) the feedback indicates only whether an answer is right

(2) the feedback supplies the correct answer

(3) the feedback is more extensive, offering a worked solution, for example.

Students find feedback that indicates only whether an answer is right to be extremely frustrating (Pascarella, 2002; Scott & Stone, 1999). Furthermore, such feedback appears to have little educational benefit. One meta-analysis (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991) found that when students were told only whether an answer was right, feedback had almost no effect on achievement (average effect size = -0.08). In a similar vein, Black and Wiliam (1998b) found that when students were given only marks or grades, they did not benefit from the feedback. However, Bangert-Drowns, Kulik, Kulik et al. (1991) found that when the students were guided to or given the correct answer, the average effect size was significantly higher at 0.31. This figure included some studies in which pre-search availability was not controlled. Thirty studies were identified which both controlled this and supplied corrective feedback (such as the correct answer or some explanation). These had an average effect size of 0.58. One conclusion drawn by Bangert-Drowns, Kulik, Kulik et al. (1991) and also voiced by Crooks (1988) and Kulhavy (1977) is that the primary benefit of feedback is error correction. Kulhavy (1977) states that the chief instructional significance of feedback is to correct errors, rather than to strengthen correct responses. This is in line with his response perseveration theory, which states that a students' response to test-like items tends to naturally persevere. He also argues that the effectiveness of error correction depends on student expectation. In particular, when students discover that an answer that they had much faith in is actually incorrect, they are much more likely to study the feedback than if they had been originally unsure of their answer: "High confidence correct answers yield the shortest feedback study times, high confidence errors yield the longest time, and low confidence answers fall somewhere in between" (p. 226). Crooks (1988), in echoing Kulhavy (1977), notes that there is a particularly great learning benefit when an incorrect response has been made by a student with a high degree of confidence in her or his response.

From the above it is clear that, after a student has attempted a question, giving him or her the right answer is of benefit. Perhaps it is also of benefit to offer further information, such as the working where some calculation is involved. For example, Charman (1999) suggests that more extensive feedback should be given even for a

correct response, since it is possible that the correct answer could have been arrived at by the wrong reasoning, or even by chance. And, in the case that both reasoning and answer were correct, he argues that feedback will reinforce this. This does sound plausible. However, as Kulhavy (1977) remarks:

> common sense suggests that increments in the complexity of feedback should act to magnify learning yields. Intuitively, the more information a student has about his response, the better he should understand why he made it. Unfortunately, the literature provides no decisive evidence that increases in complexity yield parallel increases in what is learned, and research on this point yields inconsistent results. (p. 212)

Other meta-analyses (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Schimmel, 1983) have produced similar results. Although it was expected that the kind or amount of information given in the feedback would have a significant effect, this was found to be not the case. It seems that the students are mostly just interested in the right answer and do not mindfully attend to more detailed explanations of correct or incorrect responses.

Gordijn and Nijhof (2002) investigated a computer-based instructional feedback system installed in 60 high schools in the Netherlands. A pilot study involved 149 students from 12 high schools. 537 students from 42 high schools were involved in the main study. The studies specifically investigated differences in feedback complexity, the main question being whether the amount of feedback information influenced learning gain and post-test performance. The researchers concluded that 'knowledge of correct response' feedback is just as effective as more complex feedback. The researchers were surprised by this result and suggested that the "students were inclined to memorise the correct answers rather than try to understand the concept to be learned" (p. 193).

In some situations more extensive feedback does pay dividends. This is likely to occur when the students have invested some time and effort in solving a problem but have reached an impasse. In this case, feedback that is relevant to breaking through the impasse should prove valuable. Day and Cordon (1993) found that students given help only as it was needed performed better than those students given a complete solution as soon as they got stuck. The former group were also better at applying their knowledge to similar, or only slightly related, tasks. Similarly, Scott and Stone

(1998, 1999) describe a CAL system at the University of Western Australia in which each problem is designed to 'trap' a common misunderstanding. When a student submits an answer that indicates he or she holds a particular misunderstanding, diagnostic feedback relevant to that misunderstanding is given. Each problem can also be discussed by students in a forum, offering another source of informal feedback. The evidence was that achievement standards were lifting after this system was implemented. However, feedback containing generic hints does not appear to meet with such success. One survey (Pascarella, 2002) of 500 university students enrolled in a physics course and using a computer-assisted-learning software package that supplied such hints revealed that 7 out of 8 students considered the hints to be useless. It was concluded that "time needs to be taken to program in hints that are specific to the probable errors that students are making on each problem" (p. 204). However, the system must not release the specific hints too readily. The students need to work on a problem and be ready for the hint first. This is difficult to implement in software. One approach utilises a hint or solution button for each problem that becomes visible only after a certain number of attempts have been made on that problem (Scott & Stone, 1999).

More extensive feedback may also pay dividends in terms of reduced instructor time if the feedback can be provided automatically. In the Department of Electronic and Electrical Engineering at the University of Northumbria, voluntary self-tests were provided to students via a software package. It was reported (Sambell, Sambell, & Sexton, 1999) that students tended to value the tests because they provided elaborate feedback and support materials that enabled them to work out for themselves where they were going wrong. As a result of this they no longer needed to constantly seek out tutors if they got into difficulties. Scott and Stone (1998) also reported that their CAL system had reduced instructor time to one-third its former value with no reduction in student competence.

3.6 Timing of feedback

There are two dimensions of the timing of feedback that should be considered. The first concerns the timing of the feedback in the semester. The second concerns the timing of the feedback relative to the task to which it pertains.

If feedback is provided to students with the purpose of improving their learning, then it surely follows that it must be provided while there is still adequate time for them to make appropriate changes in the way they go about their learning (Brown & Knight, 1994; Campion & Lord, 1982). 'The earlier the better' would seem to apply here. However, when it comes to implementing this ideal in practice, two difficulties emerge. First, a reasonable amount of material must be covered before it makes sense to give students a test. Second, the time taken to mark a test and return the results to the students is quite considerable. In the university context, providing a class of 500 or even 200 students with a test is a major undertaking, involving the booking of several rooms, the organising of a team of invigilators and another team of markers, and so on. Faced with these considerations, few university instructors would schedule a test before four weeks of a twelve week semester had passed. It has also been my observation that, in the subject area of electrical engineering, very few instructors set assignments or other similar unsupervised tasks that lead to students receiving feedback on their endeavours. Typically, the only formal feedback provided to students is produced by two tests and the final examination. Tutorials, which might have provided the opportunity for informal feedback, are also rare in courses from the second year onwards. Therefore, it is likely that the first substantial feedback will be from a test that takes place at least one-third of the way into the semester. If the test is a traditional one that requires human marking then it is likely to take at least two weeks to return the results to the students. The analysis above shows that students are unlikely to receive their first substantial feedback in a 12-week course before the first half of it has passed. This may not be a problem for students who do have a realistic concept of how they are progressing in the course. However, students who had a rosy view of their progress may not have enough time after receiving their test results to make up for lost ground. This second group of students would have been much better served by earlier feedback. Unfortunately, in large classes traditional testing methods are frequently too cumbersome to provide such feedback (Fernandez, John, & Netherwood, 2001). Another barrier to providing early feedback is the cost of providing feedback to students (Gibbs, 1999). For example, having three tests instead of two would provide the opportunity to deliver feedback both earlier in the semester and more frequently; however this may not be practicable in terms of financial and staff resources. These and other considerations

have led some instructors (O'Reilly, Alexander, Sweeney, & McAllister, 2002; Sambell, Sambell, & Sexton, 1999) to adopt computer-assisted assessment, and this topic is the subject of the following chapter.

The second dimension of the timing of feedback is now examined: the timing of the feedback relative to the task to which it pertains. There are many voices in the literature that speak in favour of feedback given promptly after the task: "Rapid feedback must be available to students" (Scott & Stone, 1998, p. 7), "Students... benefit from obtaining quick feedback" (Zakrzewski & Bull, 1998, p. 149), "educational research has consistently highlighted the pedagogical value of providing detailed, timely feedback to students" (Sambell et al., 1999, p. 179), "feedback should take place while it is still clearly relevant. This usually implies that it should be provided soon after a task is completed, and that the student should be given opportunities subsequently to demonstrate learning from the feedback" (Crooks, 1988, p. 468). If the feedback is not prompt then it is unlikely to be relevant because the students will be focused on new tasks. They may even have finished the course:

> three weeks after submitting an assignment, students have moved on to another topic and are tackling another assignment for another course. They may not care about anything except the mark and may not even read feedback which has been carefully and expensively provided. It is common on courses lasting a single term or even a semester for coursework submitted two-thirds of the way through the course not to be returned until after exams. This is next to useless for the purpose of guiding and improving learning. (Gibbs, 1999, p. 46)

Pascarella (2002) found that feedback delayed by just one week was often seen as being of little use.

(Kulik & Kulik, 1988) conducted a meta-analysis of 53 separate studies on feedback timing and human verbal learning. They found a variety of results, with some studies actually coming out in favour of delayed feedback: "Applied studies using actual classroom quizzes and real learning materials have usually found immediate feedback to be more effective than delayed. Experimental studies of acquisition of test content have usually produced the opposite result" (p. 79). On closer examination of the studies they found that the advantages of delayed feedback only seemed to be apparent in somewhat contrived situations. In more natural situations, and where the tasks appeared to present a reasonable level of cognitive demand,

immediate feedback was superior. They concluded that: “delayed feedback appears to help learning only in special experimental situations and that, more typically, to delay feedback is to hinder learning” (p. 94).

Students also appreciate prompt feedback (Fernandez, John, & Netherwood, 2001; Pascarella, 2002; Ricketts & Wilks, 2002). Freilich (1987) conducted an extensive survey of 213 students and 23 teachers in the area of electrical and electronic engineering to determine what factors help students learn better. Two items were highly rated by both students and teachers. One was prompt feedback (the other was study problems on new material). Mehta and Schlecht (1998) carried out a survey of students in 12 classes of varying sizes. 94% of the students indicated that getting quick feedback on their achievement was more important than class size, in itself often regarded as a major educational issue.

On the other hand, Wise and Plake (1989) found that immediate feedback may contribute to students’ test anxiety. Somewhat in opposition to this, however, is the finding by Light and Cox (2001) that students who are rather anxious benefit more than most from frequent feedback. Perhaps Wise and Plake found as they did because, at the time, computer-assisted assessment was not so prevalent and therefore immediate feedback was somewhat unusual. It has also been suggested (Jacob & Chase, 1992) that immediate feedback is problematical in tests when it is given on an item-by-item basis. Certainly, a student who received negative feedback about some previously-answered questions in a test could well become anxious about attempting the rest of the test: it would make good sense to delay the feedback until after the test was completed.

In reality, there are two requirements for the provision of prompt feedback. The first is that the feedback to the students takes place soon after the assessment. The second requirement is that the assessment takes place soon after the related learning. This requires frequent assessment, however informal, followed by prompt feedback. In the school situation, relatively small class sizes make frequent informal feedback an achievable goal for which to aim (Kulhavy, 1977). In the university situation, with its prevalence of large classes, a high frequency of prompt feedback is difficult to implement. Unless money is plentiful (an unlikely scenario!), implementations

typically involve some sort of peer marking (Gibbs, 1999; Mehta & Schlecht, 1998) or the use of technology such as computers (Chen, 2004; Gretes & Green, 2000; Merceron & Yacef, 2003; Scott & Stone, 1999; Thoennessen, Kashy, Tsai, & Davis, 1997) or even cell phones (Lian & Hu, 2005). It may be argued that some of these solutions have an advantage over the rest in that they can also provide the instructors with prompt and regular feedback on student progress. This is perhaps where hiring a team of graduate student markers or making use of peer marking is at a disadvantage: it is not so easy for instructors to gather information about student progress in these cases. Computer software packages can provide regular and prompt feedback to both students and instructors. Chapter 4 will address this issue.

3.7 Summary

As the purpose of formative assessment is to inform students and teachers of learning progress, feedback is the very essence of formative assessment. To be most effective, such feedback should come soon after the related learning. Thus, two requirements must be met: frequent assessment and prompt feedback. CBA packages such as OASIS provide a way to meet these requirements.

Generally, feedback motivates and improves learning, particularly when it focuses attention on student progress towards mastering educational objectives. It is less effective in promoting learning when it awards grades and emphasises egos. OASIS is generally employed for student practice and low-stakes assignments, some of which come very early in the semester. Most students score very highly on these assignments; perhaps this de-emphasises their grading role. Surveys and interviews suggest that students do see OASIS as a learning tool.

While most students appreciate hints and model answers, the research literature suggests that extensive feedback does not have any significant advantage over feedback that simply marks an answer right or wrong and supplies the correct answer. Instructors feel that, if OASIS did provide extensive feedback, it would be difficult to stop students during practice accessing and memorising the solution methods. Research shows that such pre-search availability leads to reduced learning.

Chapter 4. Review of the Related Literature: Computer Assisted Learning and Assessment

4.1 Purpose and outline of chapter

This chapter reviews the related literature for CAL and CBA, thus continuing the literature review that began in Chapter 2 with the topic of assessment and concludes in Chapter 5 with the topic of action research methodology. Some of the main reasons advanced for the adoption of CAL and CBA are presented in Section 4.2. This Section explores the extent and nature of the adoption, and reports on the reactions, both positive and negative, of staff and students who are using CAL and CBA. In order to indicate the range of possibilities, Section 4.3 describes and discusses three implementations of CBA: the university-wide implementation of a popular commercial software package at the University of Luton, the small-scale implementation of a home-grown software package at the University of Western Australia, and the implementation of an open-source software package across a network of institutions. The literature related to computer anxiety is reviewed in Section 4.4, while that related to the equivalence of PPA and CBA is reviewed in Section 4.5. The chapter concludes with a summary, presented in Section 4.6.

4.2 Introduction

As previously detailed in Chapter 1, instructors are facing larger classes and increased workloads. Much of this increased workload is related to assessment. Thus the partial automation of assessment becomes an attractive option. As the cost of computers decreases dramatically, such an option becomes more realistic. While, perhaps, the main reason for the adoption of CAA (Computer-assisted assessment) is the reduction of marking loads, a number of further reasons may be advanced for its adoption. CAA enables institutions to increase the frequency of assessment. Such an increase could be implemented in an attempt to motivate students to practise skills in order to improve their performance. Other reasons suggested (Bull & McKenna, 2004) for the adoption of CAA include: to broaden the range of knowledge and skills assessed, to increase the promptness and quality of feedback to students and lecturers, to extend the range of assessment methods, to increase objectivity and

consistency, and to aid administrative efficiency. Administrative efficiency is certainly increased by the deployment of CAA in those situations where universities require test and examination results to be submitted electronically: in such cases the scores can be transferred directly from the software package used for assessment to the software package used for administration.

One survey carried out in 1999 in the engineering sector in the UK that elicited responses from 59 departments in 82 institutions indicated that over 80% of institutions had invested in dedicated support for the use of technology in learning and teaching (Rothberg, Lamb, & Wallace, 2001). Another survey, also conducted in 1999, that elicited 754 responses from the higher-education sector indicated over half of the respondents in the engineering sector were currently using CAA (Bull & Collins, 2002). This latter figure was significantly higher than the figure of 37% determined for the higher-education sector in general. In engineering, and indeed in higher-education in general, the use of CAA was found to be concentrated in the first two years of the programme. Bull and Collins (2002) felt this indicated that CAA was considered to be most effective at lower levels. In their analysis, they also noted that the figures did "seem to indicate that CAA in the engineering sector is predominantly used for skill-building exercises, or practice purposes, with self-assessment, and formative tests reported as being the most common form of CAA" (p. 95). Only one-third of CAA was classified as summative in engineering, a significantly lower figure than that indicated for the higher-education sector in general. However, even summative CAA was only very rarely reported as being carried out under examination conditions (Rothberg, Lamb, & Wallace, 2001).

Multiple-choice questions were found to be the most popular question type in general, and the engineering sector was no different: "The engineering sector is shown to be heavily dependent on multiple-choice question types and text numeric match questions, which generally tend to be of a simplistic design" (Bull & Collins, 2002, p. 96). Instructors may have felt that such questions were more suitable for assessing learning outcomes at lower year levels. If so, this may explain why CAA was found to be employed mainly for lower levels and also for tests that did not make a major contribution to overall course marks. It is noted that more advanced question types, enabling the assessment of higher-order cognitive skills, do exist.

However, "it seems that the full range of question types is not being fully exploited within the engineering sector at present" (Bull & Collins, 2002, p. 98).

When asked to give their perception of the attitudes held by users towards CAA on a Likert scale, ranging from 1 (very negative) to 5 (very positive), almost half of the respondents opted for 3 (neutral), indicating that there are both perceived advantages and disadvantages in using CAA. Table 4.1, based on data from Bull and Collins, (2002) shows the full results. The respondents included both users and non-users of CAA.

Table 4.1
Perceptions of attitudes towards CAA

Category	Percentage of Respondents
Very negative	3
Negative	19
Neutral	46
Positive	31
Very positive	2

When CAA users were asked to identify the advantages of using CAA, the most frequently-mentioned advantages were the time saved in marking, the promptness of feedback, and the ability to process results electronically. CAA was also generally perceived as offering a fairer marking system and allowing a greater breadth of assessment (Bull & Collins, 2002). The following were given as examples of the positive comments made by engineering respondents: 'Common experience for all cohorts', 'Improves student learning', 'Students like it' (p. 97). In two surveys reported by Rothberg, Lamb and Wallace (2001), the most commonly expressed positive comment was that CAL "aids the learning process" (p. 503), with 84% of UK respondents and 82% of Australian respondents expressing such a comment.

Disadvantages mentioned included the time taken to prepare questions, technical problems in computerising tests, lack of staff development and training, worries about security, and limitations of question types (Bull & Collins, 2002). The following were given as examples of negative comments made by academics in the engineering sector: 'Poor support currently available', 'Limited scope of question, no

discursive solutions', 'Difficult to avoid trivialisation of complex multi-stage questions' (p. 97). In two surveys reported by Rothberg, Lamb and Wallace (2001), the most commonly expressed negative comment was that "CAL packages produced by others do not meet my needs" (p. 503), with 26% of UK respondents and 27% of Australian respondents expressing such a comment. Very few commercial packages were found to be used for engineering in more than one institution, supporting the notion expressed above that academics do have their own individual requirements of software packages. Further, academics in the engineering sector were found to prepare their own packages as often as selecting off-the-shelf packages (Rothberg, Lamb, & Wallace, 2001).

Some implementations of CAA are described in the following section. These descriptions highlight the main features and issues of CAA implementation in tertiary education.

4.3 Illustrative examples of CBA

In this section, three examples of CBA are described. In providing these descriptions, the intention is to show the range of possibilities, in terms of both the software and the extent of implementation. Software packages, which may be commercial, private or open source, may be developed by a single individual or a team. Implementations may be restricted to one department, or may cover an entire university, or may even involve a number of universities that are able to work together in some way for the benefit of all. The first example is the university-wide implementation of "Question Mark Designer", a popular commercial software package, at the University of Luton. The second example, a home-grown software package at the University of Western Australia (UWA), is a small-scale but very effective implementation, essentially involving just one department and two academics (one writing the software and the other writing the problems). The third example is the implementation of LON-CAPA, an open-source software package, across a network of universities, community colleges and high schools. In this implementation, software development is not confined to one department or even one university. Further, course and assessment materials can be shared among all participating educational institutions.

4.3.1 "Question Mark Designer" for Windows at the University of Luton

"Question Mark Designer" is a well-known commercial software package and it has been implemented on a university-wide basis at the University of Luton. An optical mark-reading system was considered but rejected on the grounds that such systems are error-prone and lack flexibility. The idea of developing an in-house software package was also considered but then rejected as it was felt that the costs would have been too high. The strategic implementation of CAA was seen as a way to "maintain the quality of student learning in an environment of rising student numbers and a decreasing unit of resource" (Zakrzewski & Bull, 1998, p. 142). Indeed, "successful computer-based assessment systems enhance the quality of the learning experience for students and are not there just to demonstrate the value of sophisticated technology" (Zakrzewski, 1999, p. 171). More specifically, it was anticipated that reduced marking loads would enable staff to better pursue other areas of academic interest, and that fast feedback would enable academic staff to plan more effectively. Also it was felt that in some cases objective testing would be the most suitable way to test specific knowledge and skills within a particular part of a course (Zakrzewski, 1999).

The system was originally trialed in June 1994, when a group of 150 first-year students sat a computer-based examination. Since only a very small number of personal computers was available, three similar multiple-choice tests were constructed. By June 1995, a much expanded network and a more structured approach to timetabling had eliminated the need to design parallel examinations. Over the next few years, usage of the system increased steadily as more subject areas and more students were involved. The number of question-types being used increased too. In the year ending June 1997, there were 9000 student sittings of examinations. This academic year was also the first in which a significant number of students used the system for formative assessment (1000 student sittings) (Zakrzewski & Bull, 1998). In the academic year 1998–99 there were over 10000 student sittings of summative assessments and 3000 student sittings of formative assessments (Zakrzewski & Steven, 2000).

The use of CAA started with a pilot study in the Department of Psychology and grew steadily as an increasing number of departments implemented CAA somewhat

independently, often using their own departmental resources. In January 1998, the University of Luton took the decision to centralize the operation of CAA. The Learning Resources Centre is the primary location for CAA, containing 200 workstations in one central complex (Zakrzewski, 1999; Zakrzewski & Steven, 2000) with an additional 200 workstations being available in the local faculty areas (Zakrzewski & Bull, 1998). An examination sat by more than 200 students may involve two sittings in the Learning Resources Centre. A typical examination is one hour long and consists of 60 to 80 questions.

Work-station failure was seen as one significant risk. To cover the possibility of work-station failure, 5% of workstations are set aside as a back-up. In the event of such a failure in the first 15 minutes of an examination, the affected student can simply log on to a new machine. In the event of a work-station failure after the first 15 minutes, the student completes the examination on paper; a few hard copies of the paper being available for this purpose. Back-up machines are not used to cover work-station failure beyond the first 15 minutes of examinations as such use could interfere with a following scheduled examination session. Careful checks are conducted 15 minutes before the examination starts and again five minutes after it has started to ensure that answer files are being stored and updated.

Student anxiety in the face of a new method of assessment was also seen as a significant risk: "The high level of anxiety was recognized from the very beginning" (Zakrzewski & Bull, 1998, p. 149).

> Introducing a new method of assessment causes student anxiety and must be viewed as a pedagogic risk. Strategies must therefore be deployed to eradicate this risk. Part of the strategy to alleviate student anxiety is to introduce sample questions on the network before the examination starts. These sample questions, perhaps 7 or 8 questions, will be available to students on an open access basis and would serve a dual purpose, first to introduce the students to the differing question types and second to enable students to get used to the technology. (Zakkrezewski & Steven, 2000, p. 208)

Question and answer sessions were also provided during tutorials prior to summative assessment. These sessions, together with the provision of sample questions, did appear to reduce student anxiety to a most acceptable level:

> Any student anxiety had been overcome by previously using the system in a formative way... The evaluation of the initial trial group at the University of

> Luton suggested that computer-based examinations may enhance student performance in terms of speed of response and a reduction in perceived examination stress. The atmosphere within the computer examination rooms was most unlike the normal examination environment which has an almost palpable air of tension, worry and quiet desperation. Within the computer environment the general impression was one of calm concentration and activity. (Zakrzewski & Bull, 1998, p. 149)

Another perceived significant area of risk centred on the academic staff, some of whom might be unfamiliar with CAA. There are two issues here, software issues and question-design issues. Covering the first, academic staff who use computerised assessments are supported by a central unit which takes care of all software issues. Thus, academic staff are not required to use the software at all, and are free to concentrate on question design. In this they are given support and documentation on writing suitable test questions. Further, all staff who wish to implement CAA are required to attend a staff development programme that describes such assessment and explores the pedagogic issues surrounding the construction and use of objective tests. This programme is provided by the Staff Development Unit at the University of Luton and forms an integral part of the university's staff development programme.

The results of this implementation appear to have been most favourable. Both staff and students appreciated the quick feedback. The rapid feedback to staff enabled timely responses to any potential problem areas. Though, initially, there was some staff concern over the requirement to develop between 60 and 80 suitable questions for each examination, staff training, staff sub-division of tasks, and use of question banks, have all helped reduce this concern. In fact, staff work-loads associated with assessment have been reduced: "automatic marking, no second marking and comprehensive statistical analysis release academic staff to pursue other areas of interest" (Zakrzewski & Bull, 1998, p. 149). Once a staff member had implemented his or her first computer-based examination , that staff member was likely to make further time savings when implementing subsequent computer-based examinations. The indications were that there was "an average saving in academic staff time of 50% for first-time computer-based examinations... A further saving of 50% was achieved for subsequent updates" (Zakrzewski & Bull, 1998, p. 150).

The benefits associated with rapid feedback and reduced examination stress are significant, but perhaps the most significant benefit reported at the University of Luton for the students involved with CAA is enhanced learning, as measured by examination results. The perceived improvement in student learning is attributed to formative assessment, something certainly made far more manageable in terms of staff workloads by the use of computers. For example, three biology modules had previously been assessed through optically-marked multiple-choice examinations. For 1997, the assessment, still using the same bank of questions, became computer-based. Formative assessment was provided in two modules, each of which produced a 1.6 grade point increase, while the module which had no formative assessment produced no change: "Overall the first-time pass rates showed an average increase of 7.5%" (Zakrzewski & Bull, 1998, p. 150).

Overall, the implementation of CAA at the University of Luton would appear to have been well planned with a high level of attention paid to dealing with potential risk factors. Staff workloads were reduced by the automation of many aspects of the assessment, while prompt and comprehensive feedback to staff and students has increased satisfaction, responsiveness, and the level of information available to those involved. Finally, the ability of CAA to readily deliver formative assessment has enabled student learning to be significantly enhanced. However, it is clear that not all staff utilise CAA at the University of Luton. Though no reasons for their decisions are given in the published literature, these reasons may well include resistance to change, and some apprehension about having to produce 60 or 80 questions of a new style for each examination.

The centralised operation described above is similar in some respects to the CML testing system in operation at Curtin University of Technology. This mainframe-based system, utilised predominately for first-year classes across a range of subjects including economics, education, human biology, instrumentation, medical imaging and psychology, is used by several thousand students who take, on average, two or more tests per semester. Tests can be generated and marked by the software, with the results subsequently analysed and stored. Available feedback covers both student performance and the effectiveness of individual questions in the question-banks. The CML system is routinely used for both formative and summative purposes. As

outlined in Section 2.3, an analysis of the data collected by the system clearly shows the value of formative assessment (Sly, 1999; Sly & Rennie, 1999).

4.3.2 "Dynamic Teaching Solutions" at the University of Western Australia

Dynamic Teaching Solutions, a home-grown software package delivered via the Web, was produced as a replacement for the traditional tutorial system in a first-year dynamics course in the Department of Mechanical Engineering at UWA. The move to a computer-based tutorial system was motivated by continued low attendance at tutorials, together with a failure of a significant number of students to visualise basic dynamics concepts. The software package evolved over a number of years, starting in 1995 as a network-based application, and later metamorphosing into a web-based application, with one academic creating and maintaining the software while a second academic was producing the problem sets (Scott & Stone, 1998, 1999). Since these problem sets are replacing tutorials, it is clear that the intention is to provide formative rather than summative assessment. Each set consists of 10 to 20 problems, the majority being relatively easy 'lead-up' problems that enable students to develop their skills before they reach the harder assessed problems. Each student must complete all 'lead-up' problems before she or he is able to attempt the two to four assessed problems. Each problem set must be completed before a stated deadline. Different students have different deadlines for each problem set, the deadlines being chosen randomly in order to spread the load on computing resources and to ensure fairness. Students may attempt all questions, including assessed questions, more than once; however, marks are lost every time an incorrect answer is entered for an assessed question. Altogether, the assessed questions count for 20% of the course mark while examinations count for the remaining 80%.

Ideally, each problem targets one or more misconceptions, with most students obtaining either the correct answer or one of a few incorrect answers, each of which corresponds to a particular misconception. However, some questions are found to produce a high arithmetic error rate, and these questions may often fail to identify misconceptions. When an expected incorrect answer is detected, diagnostic feedback is given to the student. For each problem, a full solution is available. In the case of assessed problems these solutions are not available until after the deadline. For each

non-assessed problem, students can access the solution after a set number of attempts.

The current tutoring system provides each student with different numerical values; hence the values for correct and incorrect answers are also different. This individualisation of problems prevents students from helping each other by simply stating answers to problems. Students may, of course, still collaborate by explaining to each other how to do various problems. Such collaboration was not seen as a negative feature of the system but was in fact encouraged: "The environment must foster student-student collaboration; this must occur naturally and not be arranged in extrinsic or arbitrary ways. Students must be kept together in both space and time" (Scott & Stone, 1998, p. 7). Of course, once the computer-based tutorial system moved out of a room full of computers and on to the Web, such collaboration became less natural, with a significant number of students working on problems from home:

> Actually we were somewhat apprehensive about the idea of students working at home alone on the problems, since we are strong believers in the value of collaboration in problem-solving and learning. As soon as students realised that the work could be done from home we had many enquiries about the hardware and software needed to do so. (Scott & Stone, 1998, p. 15)

However, two hours per week in the computer laboratory are timetabled in a deliberate move to foster student-student collaboration. Tutors are present during these timetabled sessions; however, the laboratories are also open and free at other times for student use.

In order to promote good practice, students are required to record all their working for tutorial problems in a log book. This log book is checked when students ask a staff member for help: students are turned away unless they are able to show clear working in their logbooks. There is also a discussion forum for each tutorial problem: such forums provide a useful way for students to help and collaborate with each other. The staff can rapidly judge which problems are causing difficulties, and the forums also provide highly efficient vehicles for answering student queries in a large class:

> The volume of traffic on the forum was a good indicator of the difficulties that the class was having. It proved to be a remarkably efficient way of dealing with

> basic questions. Such questions only had to be answered once for the whole class. (Scott & Stone, 1999, p. 107)

A monitoring system enables staff to see the progress of each student through the problem set. The system graphically shows how many problems each student has successfully completed and gives an indication of how many attempts were needed to master each problem: "The value of the monitoring system cannot be over-stressed. Students who are falling behind are immediately detected. Such students can be found and help provided" (Scott & Stone, 1999, p. 105). All student problem-solving activity is logged, providing a rich source of data for future analyses and more considered responses to student performances. In the four years since the introduction of the tutoring system in 1995, data on 250 000 student solutions have been acquired.

Students have received the system enthusiastically since its inception, even though, in its first year, 1995, there were frequent computer crashes which entailed the students having to redo some of their work. In spite of data losses and delays while the system was rebooted, the students that year were "helpful and forgiving, and they asked that work on the tutorial system should continue... it was not long before they were asking why other courses did not use the same computer tutorial system" (Scott & Stone, 1999, p. 104). Through the next year, by contrast, the system ran most reliably: "1996 was truly wonderful. The computer system... was fully in place and fully debugged... the students learned extraordinarily well and were especially happy, and yet the effort required to teach and to maintain the system was minimal" (Scott & Stone, 1999, p. 104).

The pattern of 1995-6 repeated itself in 1997-8 with the introduction of the Web-based version. The first year of the Web-based version, 1997, saw frequent crashes; "in the first weeks of operation it failed or crashed every few hours; towards the end of the semester it was running for several days without failure" (Scott & Stone, 1998, p. 15). Students who were keeping log-books as required "could in theory simply retype lost work and re-submit it, but there were still some frustrating delays" (Scott & Stone, 1998, p. 15). By 1998 the new system was relatively free of bugs.

Students also appreciated the way in which the software would identify common misconceptions and deliver appropriate feedback. This appreciation became clearer when the Department of Civil Engineering at UWA adopted a similar tutorial system, but one without diagnostics. This system simply gave students a right/wrong response: "Student surveys showed that the lack of diagnostics was a significant loss to the students" (Scott & Stone, 1999, p. 107). The question of how much feedback is appropriate was discussed in Section 3.4; the feedback provided by Dynamic Teaching Solutions was very specific. Such specificity is very important because feedback beyond knowledge of the correct answer is of little help if it is untargeted, as researchers such as Pascarella (2002) have found:

> The hints that were programmed into some CAPA problems were considered by many of the interview subjects (7 out of 8) to be useless. (p. 202)
> time needs to be taken to program in hints that are specific to the probable errors that students are making on each problem. (p. 204)

OASIS does not offer feedback beyond knowledge of the correct result and a number of students have requested more elaborate feedback, from hints to full solutions. However, a few students have spoken against the provision of more elaborate feedback. Subsection 8.2.3 explores this aspect of OASIS in more depth.

Students judged that the computer-based tutorial system at UWA had enhanced their learning. The implementation of the system also reduced staff workload:

> It has been observed that, in the years that this teaching method has been used, student competence (as measured by formal written examinations) is comparable to that of previous years, that student satisfaction is high, and that staff teaching time has been reduced to about one third of its former value. (Scott & Stone, 1998, p. 9)

With the introduction of a new learning method that enhances achievement levels, there is the danger that the difficulty of the tests and examinations set will be increased without the examiners being fully aware of such an increase. Certainly, some increase in examination standards would appear to be taking place in the year-one dynamics course:

> The level of competence demonstrated by students in the June 1997 examination was definitely comparable to that in previous years. In fact we believe that our examinations have become progressively harder, and that we are suffering from creeping standards, and yet we also observe a gradual increase in the mean mark. (Scott & Stone, 1998, p. 15)

Unsupervised student work that is assessed always suffers from the problem that students can rely on the efforts of others rather than make the effort themselves to master the material. This problem was decreased in the move to individualised questions that came with the Web-based tutorial system. However, the original networked tutorial system did not have individualised questions and was particularly vulnerable in this respect:

> Under the heading 'Do you have any suggestions for change?' one student wrote: 'Have some way of making me do [the problems] without looking at other peoples' answers. Often, when the deadline is close, it is easier to cheat to satisfy the requirement without learning anything.'
> (Scott & Stone, 1998, p. 16)

Scott and Stone are highly aware that implementations of computer-based systems do not come without social risks. Their care in maintaining student-student collaboration has already been noted. Some students may well see computers as too impersonal and may wish for more interactions with humans. This problem is compounded when tutorials, discussions, simulations, and perhaps even lecture notes, are all delivered via the Web. Staff must make strenuous efforts to counteract the impersonal quality introduced into courses by a heavy reliance on computers: "Teaching methods that use computers can alienate a group of students and this effect must be counteracted by particular attention to the social side of learning. We believe that a computer-based teaching method can only succeed where staff are energetic and enthusiastic" (Scott & Stone, 1998, p. 7).

Scott and Stone also describe a recently-developed new type of problem: "Students are shown a problem and a solution. They are then required to indicate any errors in the solution... It is of course possible to give a correct solution and to see if students consider errors to be present" (Scott & Stone, 1999, p. 108).

4.3.3 LON-CAPA at Michigan State University and Simon Fraser University

LON-CAPA (Learning Online Network with Computer-Assisted Personalised Approach) is a free software package. The code base is open source, and the software platform is free of licensing fees. It is used in approximately 24 universities, together with some high schools and community colleges, almost all of which are located in the United States. The software originated at MSU where it is currently used by around one quarter of the student population. The second-largest group of users is at

Simon Fraser University. CAPA itself was piloted in 1992 with a physics class of 92 students. In its initial form CAPA was not Web-based.

> CAPA provides students with personalized problem sets, quizzes, and exams. Students are given instant feedback and hints via the Internet and may correct errors without penalty until the assignment due date. The system records the students' participation and performance and the records are available online to both the instructor and the individual student. (LON-CAPA, 2005)

A second project, LectureOnline, was initiated in 1997 with 770 physics students.

> With only the web-browser as interface, LectureOnline enables instructors to seamlessly put together a presentation of material gleaned from all over the Internet, and to create different types of individualized online homework. Grading, communication, groupwork and enrollment are also handled by the system. (LON-CAPA, 2005)

The groups associated with these two software packages joined forces in 1999 to produce LearningOnline Network with CAPA, or LON-CAPA. This new package, first trailed in 2001, provides a superset of the functionalities offered by the two original software packages.

> LON-CAPA is a distributed open-source Learning Content Management and Assessment System that provides instructors with a common, scalable platform to assist in several aspects of teaching a course, from lecture preparation, communication among faculty and students, calendar keeping and announcements, to administration of homework assignments and exams. It also enables instructors to create educational materials and to share such learning resources with colleagues across institutions in a simple and efficient manner. (Kortemeyer et al., 2003)

The "top five value-adding features of LON-CAPA" (LON-CAPA, 2005) are seen as being: individualised assessment, immediate feedback, open-source freeware, one source - multiple target, and content sharing. Each of these is now looked at in turn.

Individualised assessment may entail scrambling the multiple-choice options for a question so that different students receive the same options but in a different order. It is also possible to provide students with numerically different versions of the same question, or even with completely different questions. Such individualisation reduces cheating by students and makes it easy for staff to prepare different versions of the same examination. In spite of this individualization, the software package still enables academic staff to easily pinpoint misconceptions and problem areas early in the course and take corrective measures, such as just-in-time teaching.

LON-CAPA, in common with other CBA systems, takes care of the marking of student submissions, freeing staff from the drudgery of large-class marking, and produces rapid feedback which is highly valued by students and staff alike. Feedback to students can be targeted to specific misconceptions. Feedback to staff is sufficiently detailed to reveal how many attempts it took each student to get each problem correct. The version of each problem received by each student can be seen, as can all answers submitted by each student. One problem can be analysed across all students to see which options the students selected first, for example. In spite of the fact that each student has a different version of any given problem, complete item analysis is still possible.

As open-source freeware, LON-CAPA offers a number of advantages. Implementing the software involves a low monetary investment, freeing up funds for support and training. For those relatively few institutions prepared to take on the task, there is the prospect of modifying the code to suit local requirements. These requirements could be driven by campus, hardware or course considerations, or even by research considerations. For example, a researcher may wish to modify the software so as to record, display or analyse certain data in a particular way. Certainly, the code is more open than with commercial packages and can be continually improved by the wider community. With LON-CAPA, the pool of skilled users is now sufficiently large to make this continual improvement a reality, although in fact most of the contributions to source-code improvement continue to come from programmers at MSU. Perhaps this is to be expected: most users of LON-CAPA would not be sufficiently adept at programming to venture to produce improvements. Another advantage with LON-CAPA is that institutions have their own servers. With comparable proprietary systems, the server may be located outside the university. Some registration and communication problems have been reported for such arrangements, and there are privacy issues to consider too (Batchelor & Jungic, 2004).

LON-CAPA is ‘one source - multiple target’: the same assessment content can be used for online homework, PDAs, printed Scantron examinations, and online examinations. A single question can be delivered in a range of question-types. Personalised hard-copy printed examinations consisting of individualised questions

can also be produced in order to make traditional examinations even more secure (Albertelli, Kortemeyer, Sakharuk, & Kashy, 2003). The print quality of any hard copies can be varied depending on whether paper-saving or presentation quality is of more importance.

Content sharing is a highly useful feature available to users of LON-CAPA and is implemented with a high degree of flexibility with different levels of granularity being possible. For example, an instructor at one institution is able to select a single problem, an assignment, or even a whole course from any other participating institution. Text and movie clips may also be sourced from other institutions.

> An instructor from Community College A in Texas can compose a page by combining a text paragraph from University B in Detroit with a movie from College C in California and an online homework problem from Publisher D in New York. Another instructor from High School E in Canada might take that page from Community College A and combine it with other pages into a module, unit or section. Those in turn can be combined into whole course packs. (Minaei-Bidgoli, 2005, p. 15)

Content sharing takes place not only across institutions but also across disciplines. For example, an individualised problem type that originated in the area of botany went on to become extensively used in the teaching of astronomy.

Content sharing is possible because each institution using LON-CAPA provides at least one server for the hosting of LON-CAPA and these servers are constantly connected with each other through two-way TCP/IP connections, thus forming a geographically distributed network of servers. The process of content sharing is kept simple and reliable by the fact that the authoritative version of any resource always continues to reside on the author's server.

> Users can log into any server in the network. For example, an MSU user can log into a server at North Dakota State University using his MSU credentials, and find the exact same environment as on one of the on-campus servers. To faculty users, the distributed content resource pool of LON-CAPA appears as one large virtual file system, and every resource has a system-wide unique and persistent URL path, under which it can be accessed from any server in the network. As users are browsing this filesystem, they are actually transparently accessing content from servers across the network... As resources are accessed, the network provides transparent resource replication to provide faster access to the resources. (Kortemeyer et al., 2003)

The sharing of educational content across institutions has three desirable consequences. First, it helps establish cross-institutional standards and can be used to promote bench-marking. Second, it appears to be a significant motivator in encouraging academics to create quality course materials: “The potential of having an impact beyond the boundaries of one's own classroom appears to be a major incentive for the creation of such resources for many educators” (Kortemeyer et al., 2003). Authors do not appear to be held back by issues of royalties and intellectual property; of course, authorship of each resource is acknowledged on the system.

Third, the quality of the course materials is enhanced by the fact that many different academics are using them. Academics may find mistakes or suggest improvements in materials authored elsewhere: “The peer-review provided by the resource selection mechanisms, as well as the continuing evaluation by thousands of students across courses, semesters, and institutions increases resource accuracy, reliability, and relevance” (Kortemeyer et al., 2003). The knowledge that one’s resources are being widely read and used must surely be a strong incentive to take special care in producing them in the first place.

Students give the system high grades for helping them achieve study goals: “83% of the students rated CAPA as either very helpful or helpful... with only 9% giving it a negative rating” (Thoennessen & Harrison, 1996, p. 146); “A majority of students, typically 80%, consider that CAPA helps them learn and understand the course material” (D. A. Kashy, Albertelli, Kashy, & Thoennessen, 2001, p. 499).

One study (D. A. Kashy, Albertelli, Kashy, & Thoennessen, 2001) reported on the difference LON-CAPA made to the grades gained by an introductory calculus-based physics course. In the period 1992 to 1994, before LON-CAPA was implemented, the final grade distribution exhibited the traditional bell shape around a grade of 2.5, with only 20% of students receiving a grade of 3.5 or 4.0. After the move to LON-CAPA, the proportion of students earning a grade of 3.5 or 4.0 increased steadily to reach 36% in 2000. In order to protect against the possibility that a lowering of standards for the class had produced increased grades, independent evaluators were used to judge the examinations. They reported that the examinations used in the course after deployment of LON-CAPA were more challenging than those used in

earlier years. Female students were found to benefit even more than their male counterparts. Other studies have also found that LON-CAPA has lifted student performance (Thoennessen & Harrison, 1996) and reduced the number of students dropping out of courses (E. Kashy, Thoennessen, Tsai, Davis, & Wolfe, 1997).

There are a number of reasons for the increased student achievement outlined above. The first is that the prompt feedback delivered by LON-CAPA enables students to recognise, focus on and master the material they do not know. This feedback also enables staff to teach in a more informed manner. Second, being able to set firm electronic deadlines helps students keep up-to-date with their work. Frequently, instructors would set short assignments due on a weekly basis, helping students establish regular and effective work-habits. Third, students using LON-CAPA actually greatly increased their time spent on tasks: "Time students spend working on assignments and other course requirements has increased by nearly a factor of two and approaches the recommended two-hours outside of class per lecture hour" (D. A. Kashy et al., 2001, p. 499). Clearly, in order to bring about this increase in time spent on task, there was something highly motivating about LON-CAPA: possibilities include the prompt feedback and the regular short assignments and deadlines. The ability to re-do questions was also seen as a motivational factor: "Allowing multiple tries on assigned problems with no penalty is highly motivating; most students strive to get all the work done correctly" (D. A. Kashy et al., 2001, p. 500). In fact, a maximum number of attempts can be set for each question, and attempts beyond the first can be with or without mark penalty. However, in general, academic staff do allow multiple tries without penalty. Setting a maximum number of 15 repetitions is typical. Such a high number does allow students to score highly but it also encourages guessing, particularly where question items are effectively a combination of a small number of true/false questions. Some students are also tempted to try to guess the correct equations for simple numerical problems (Pascarella, 2002).

LON-CAPA also provides excellent opportunities for student-student and student-staff communication: directly by providing message forums, indirectly by freeing staff from high marking workloads: "There is a high level of interaction among students and between students and staff. A smaller teaching staff can do more for (and with) students. Reassigned staff can provide a greater level of Socratic

interaction with individual students" (D. A. Kashy et al., 2001, p. 500). Such increased interaction can motivate students as well as provide help to lift their performance.

Unfortunately, while some educators are using technology to promote learning, some students are using technology to circumvent learning. While personalised problem-sets render copying of answers irrelevant, spreadsheets and web-sites are being used by students to avoid putting in the work needed to master the material. Such avoidance does ultimately result in decreased final grades.

> Students have established interactive web sites (financed by banner ads and donations) to network with each other in an attempt to defeat the system. Whole Excel spreadsheets are going up in an attempt to reverse-engineer the individualization mechanisms – one problem at a time, and featuring '52,359 homework forum messages.'... a survey in which students after the end of a course were asked how often they used the cheat site has shown that the final grade was significantly negatively correlated to site usage. (Kortemeyer et al., 2003)

In 1999, a decrease in performance was noted in the student group in one physics course. Through 1996 to 1998, 78% of the students had achieved grades of 2.5 or above. This figure dropped to 70% in 1999 but lifted to 77% in 2000, in keeping with the years before 1999. It has been suggested that the decrease in performance in 1999 was caused by one student providing a web-site that enabled students to complete their homework assignments without having to understand the material.

> An enterprising student developed an elaborate web discussion forum where students could get answers and formulas, often with little understanding, thus defeating the goals in the design of most individualized numerical and conceptual problems. A recent analysis shows that students' use of that web site is negatively correlated with examination performance, ($r = -.35$, $p < .001$), i.e., students who used it more tended to score lower on midterm exams, quizzes, and the final... In contrast the correlation for students using the discussion site provided with the course was positive... These results were described to the students in 2000, and the warning may have been sufficient to neutralize the negative impact of the student-run web site. (D. A. Kashy et al., 2001, pp. 501-2)

In an attempt to render such web-sites and spreadsheets ineffective for those students who ignore the warnings about using them, the instructors using LON-CAPA are developing new problem formats that make copying without understanding much more difficult; random labelling on individualized graphs, for example. It may be

expected that student use of technology to defeat technology will be an ongoing issue: "The final word on this 'technology war' has not been spoken yet" (Kortemeyer et al., 2003).

Another problem sometimes encountered in the use of technology such as LON-CAPA is the lack of understanding displayed by administrators. Although administrators have, in the majority, been found to be supportive, a number of cases have been described where administrators have hindered or even blocked the deployment of such packages (E. Kashy, Thoennessen, Albertelli, & Tsai, 2000).

In spite of the problems outlined above, and the fact that workloads may be high during the initial implementation stage, instructors using LON-CAPA express satisfaction and sometimes enthusiasm with the software package.

> The level of satisfaction with the on-campus ALN [Asynchronous Learning Network] implementation is high across many disciplines and faculty who have implemented all or part of its functionality. This satisfaction comes in spite of the universal agreement among faculty that work is increased, especially initially... Positive interactions with satisfied students, by far the majority, is a big factor, as is the interactions with colleagues doing ALN with whom one can share a remarkable variety of wonderful stories. The on-campus aspect has strong appeal. One faculty who volunteered that he is now '... a convert to this technology' added that he liked that he was 'still teaching the normal way.' Another commented, 'This was the first time I had the ability to really see how students were doing in such a large course and could review that information before meeting with them.' (E. Kashy, Thoennessen, Albertelli, & Tsai, 2000, p. 239)

The increase in quality student-staff and staff-staff interactions commented on in the above is also echoed by others. The interactions are frequently centred on student learning, often a topic that previously received little attention.

> Research indicates that discussion among students and between students and staff is greatly enhanced when LON-CAPA is implemented in a course. Another interesting effect has been the level of discussion dealing with student learning which now takes place among faculty using the technology. This has led to considerable sharing of methods and applications among the faculty. This is a great step forward considering that faculty collaboration in teaching activities is often an essentially nonexistent aspect in many departments. (Kortemeyer et al., 2003)

It is interesting to note that staff do feel positively about using LON-CAPA even though one of the aims of CAA, workload reduction, does not seem to have been

universally realised. The reason for the lack of workload reduction may well be an overall increase in assessment. For example, some users of LON-CAPA set their students weekly assignments (Batchelor & Jungic, 2004). With the large classes prevalent in higher education, such weekly assignments would be impossible without automated marking. In my own department, in courses that use OASIS assignments or tests, the CBA introduced tends to be more than the PPA replaced. The reason for staff enthusiasm in spite of continuing high workloads is to be found in the nature of the work, as outlined above by E. Kashy et al. (2000) and Kortemeyer et al. (2003). Time previously spent on relatively mundane low-level activities, such as marking, is now spent on more rewarding higher-level activities, such as interacting more directly with students and colleagues, and reviewing current overall performance data.

LON-CAPA collects a wealth of information about student performance and this information may be analysed by instructors to help predict future success, identify at-risk students, and identify the sort of behaviours that lead to success. Without adequate technology, such valuable outcomes are simply not feasible. One example of such an analysis was carried out at MSU (Minaei-Bidgoli, Kashy, Kortemeyer, & Punch, 2003), where an attempt was made to predict final grades of students based on the students' behaviours through the semester. The approach involved designing, implementing and evaluating a number of pattern classifiers. Subsequently, a combination of multiple classifiers was found to produce the best classification performance, and this combination was modified using a genetic algorithm to further improve prediction accuracy. Six variables were chosen as predictors of the final result for every student. The first five variables were based on how the students performed with their regular homework assignments and were as follows: success rate (total number of correct answers), success at the first try, number of attempts before correct answer is derived, the time at which the student got the problem correct relative to the due date, and the total time spent on the problem. The sixth variable was the number of online interactions of the student both with other students and with the instructor.

In predicting whether students would gain a grade regarded as high (3.5 or 4.0), medium (2.5 or 3.0) or low (2.0 or less), the total number of correct answers was the

most important variable (Relative importance = 100%) followed by the total number of tries (58%), success on the first try (28%) and total time spent on the problems (25%). In an extension of the original investigation (Minaei-Bidgoli, 2005), in determining whether students passed or failed a course, the variable 'number of problems correct with a high number of tries (10 or more tries)' was found to be the most important variable (Relative importance = 100%) followed by 'total number of correct answers' (92%) and 'total time spent on the problems' (40%). That the variable 'number of problems got correct with a high number of tries (10 or more tries)' was found to be the most important determinator of whether students passed or failed seems remarkable. Ten or more attempts before a problem is correctly answered seems excessive and this number does lend weight to the notion that students did frequently try to guess answers (Pascarella, 2002).

The variable 'total time spent on the problems' can only be measured in terms of time logged in: if there were some way to measure the amount of time students were actually mindfully engaged with the problems rather than just logged in, then this variable would most likely have returned a much higher relative importance figure. When genetic algorithms were employed as outlined above in predicting whether students would pass or fail a given course, prediction accuracies slightly in excess of 90% were achieved across the dozen or so courses for which the predictions were attempted (Minaei-Bidgoli, 2005). For a particular physics class of 227 students, the prediction accuracy was 94%. When the same method was applied to predict which of 9 grades (0.0, 0.5, 1.0... 3.5, 4.0) each student would be awarded, the prediction accuracy was 62% (Minaei-Bidgoli, Kashy, Kortemeyer, & Punch, 2003). Such figures are really quite remarkable, and indicate that these sorts of research methods can clearly identify student study behaviours that lead to success and failure.

Armed with the sort of knowledge outlined above, instructors can intervene at an early stage and attempt to modify student behaviour for the better. In one first-year introductory physics course for engineers, the performance of the 400 enrolled students was monitored by a software tool that could predict their final grade at any desired time during the semester. This prediction was based on homework, quiz and midterm examination performances, where applicable. One third of the way through one particular semester, when the results of five homework assignments, three

quizzes and one midterm examination were available, the software tool was used to predict which of these pre-engineering students were likely to gain a grade of 1.5 or lower (a grade of 2.5 or more was needed for entry to the engineering programme). One hundred such students were identified in this way, and a second software tool was used to generate and send individual messages to these students. A typical email is quoted below.

> Dear David,
> You are enrolled in PHY184 and based on your current performance you will receive a grade of 0.0. You have solved 35.0% of the homework problems and you missed 1 of the 4 quizzes. Extrapolating from these data and from your exam 1 score of 19.5% your final percentage is projected to be 51.7% which corresponds to a grade of 0.0.
> It seems that you have some difficulties with this course. Please contact me to discuss your situation and possible improvements either by e-mail (thoennessen@nscl.msu.edu) or by phone (5-9672, ext 323) to set up an appointment.
> Regards,
> Michael Thoennessen. (Thoennessen et al., 1997)

The response to this automated mail-out was most impressive:

> 53 students replied within a week either requesting an appointment and/or explaining their poor performance and promising to improve. Almost all students were positively impressed by the detailed information given and by the personal tone of the message... They were mostly surprised that a professor in such a large class took the time to contact individual students.
> (Thoennessen et al., 1997)

The intervention appears to have been successful. In fact, the students who were predicted to gain a final grade of 1.5 and were emailed ended up with grades that were, on average, as good as those students who were predicted to gain a final grade of 2.0, and who were not emailed.

LON-CAPA is a highly sophisticated and successful package that offers a wide range of possibilities for students and instructors. One feature that it, surprisingly, did not offer until very recently is the ability for students to practise a problem over and over again with different numerical values. This is an important feature of OASIS. LON-CAPA allows students to repeat homework problems till they get them right but the numbers in numerical problems are fixed. In June 2005 I emailed some of the key users/developers of LON-CAPA about this and received the following responses:

> The best you can do is publish a problem as public. And have the students logout of lon-capa and go to the url of the problem. (Albertelli, G.)

> Another thing you can do is include the same problem [in a homework set] several times - you will have a different randomization every time. (Kortemeyer, G.)

> As regards the question of practice problems: Guy Albertelli will probably soon act on the suggestion of making this functionality part of the LON-CAPA release (version 1.99 is just out and I expect version 2.0 to soon follow). Since Guy already has the code that makes the public versions of such problems work, it could well be included in the v2.0 release. Three other major users also posted messages indicating that they too want this to be a regular feature. (Batchelor, R.)

Many instructors use LON-CAPA to deliver weekly homework assignments so perhaps they do not see this sort of repetitive practice as relevant for their students. Courses using OASIS tend to have less frequent assignments (perhaps fortnightly or monthly) and more emphasis on practice. In some cases, OASIS has been used successfully to provide practice opportunities without any assessment. For example, in 2002 OASIS was used to provide practice without assessment for approximately half of the year-one Electrical Engineering Systems course taken by all 550 year-one engineering students. OASIS was well received that year and the average student submitted answers to 100 questions.

4.4 Assessment and computer anxiety

The issue of assessment anxiety, already raised in Section 2.2, noted that the higher the assessment stakes, the greater the anxiety, which in turn may result in decreased performance. Assessment anxiety is also increased when the assessment is unfamiliar or difficult (Crooks, 1988). Perhaps because they are perceived as more difficult, less frequent and longer assessments produce more anxiety than more frequent and shorter assessments (Fulkerson & Martin, 1981). Computers are in fact well suited to the delivery of short, regular assessments or tests and are routinely used to deliver the same on a fortnightly or even weekly basis (Batchelor & Jungic, 2004; O'Reilly, Alexander, Sweeney, & McAllister, 2002; Scott & Stone, 1999). These assessments are typically low stakes and also in many cases allow students to repeat questions several times without penalty until the correct answer is achieved (Deeks, 1999; Grebenik & Rust, 2002; Thoennessen & Harrison, 1996).

The prompt feedback that is a feature of CAA can help reduce anxiety: it has been noted that a lack of feedback on performance increases anxiety (McDowell, 1995), and this lack of feedback is something students in large classes frequently contend with as they wait three or four weeks for an assessment to be marked and returned. However, it is possible for software packages to increase assessment anxiety by delivering feedback too soon (Wise & Plake, 1989): for example, there is some evidence to suggest that software packages that deliver feedback on test items prior to the submission of the completed test can increase anxiety (Jacob & Chase, 1992).

Currently, the vast majority of students have ready access to computers and the Internet (Palmer & Bray, 2001). Therefore, they are not likely to suffer computer anxiety caused by a lack of familiarity with computers themselves. However, it is possible that anxiety could be brought on by the use of a software package with which the students are unfamiliar. To counter this difficulty, it is normal practice for institutions to run practice tests prior to the actual high-stakes assessment using the same software under very similar conditions (Sly & Rennie, 1999; Zakrzewski & Bull, 1998). The implementation of practice tests has been found to reduce anxiety to an insignificant level. In fact, at the University of Luton, it was felt that examination stress was actually decreased through the use of CBA:

> The evaluation of the initial trial group at the University of Luton suggested that computer-based examinations may enhance student performance in terms of speed of response and a reduction in perceived examination stress. The atmosphere within the computer examination rooms was most unlike the normal examination environment which has an almost palpable air of tension, worry and quiet desperation. Within the computer environment the general impression was one of calm concentration and activity. Student responses to pre and post examination questionnaires were very positive. (Zakrzewski & Bull, 1998, p. 149)

Of course, it does not follow that students familiar with computers will necessarily feel no anxiety about using them. In a survey of microbiology students sitting a computer-based, multiple-choice test, it was found that 61% of the students had no 'computer phobia' while 24% had low and 15% had moderate to high phobia (G. Lee & Weerakoon, 2001). Of this group of students, 88% stated that they had a computer at home, and 82% responded that they had moderate to high experience with word processing on computers. Clearly, some students exhibiting some degree of phobia did have home computers and were familiar with some aspects of their use, at least

sufficient to master the mechanics of a multiple-choice test. The above result suggests that computer anxiety is determined not just by computer experience but also by perceptions. Backing up this notion, a study of first-year business undergraduates (Anderson, 1996) revealed that computer anxiety was related to perceived knowledge of computers, rather than actual computer experience. Anderson also found that gender was not a significant factor in predicting computer anxiety: "The results of this study do not support the contention that women in general exhibit higher levels of computer anxiety than men" (Anderson, 1996, p. 75).

Somewhat in contradiction to Anderson's findings, a meta-analysis of 36 studies of computer anxiety published between 1990 and 1996 found the following results:

> (1) female university undergraduates are generally more anxious than male undergraduates, but the strength of this relationship is not conclusive; (2) instruments measuring computer anxiety are generally reliable, but not compatible with one another; and (3) computer anxiety is inversely related to computer experience, but the strength of this relationship remains inconclusive. (Chua, Chen, & Wong, 1999, p 609)

Given that Chua, et al started with 60 studies and reduced them to 36 studies worthy of further consideration, the tentative nature of their conclusions is not encouraging. Chua, et al found the extent of the associations varied considerably between studies; for example, in 12 studies female university undergraduates displayed significantly more computer anxiety than did male under-graduates and in seven studies the difference between male and female students was found to be insignificant. The wide variation they found among the results of investigations into links between computer experience and computer anxiety is also noted in another survey of the relevant literature: "It has been assumed that computer anxiety results from a lack of familiarity with computers… but research has produced conflicting findings… the belief that computer anxiety results solely from lack of exposure… appears too simplistic" (McDonald, 2002, p. 305).

The above findings indicate that there are difficulties in reliably linking computer experience to computer anxiety. There are also difficulties in reliably linking computer anxiety to performance in computer-based tests. For example, in the previously-mentioned study of microbiology students sitting a computer-based, multiple-choice test (G. Lee & Weerakoon, 2001), neither computer experience nor

anxiety correlated significantly with test performance. A similar result was found in a study of a computer-administered mathematics placement test (Shermis & Lombard, 1998); test anxiety but not computer anxiety was a significant predictor for test performance (at the $p<0.01$ level). However, Shermis and Lombard found that, for a computer-administered reading placement test, computer anxiety but not test anxiety was a statistically significant performance predictor (at the $p<0.05$ level, but not at the $p<0.01$ level). Further, a study (Anderson, 1996) of business undergraduates undertaking an introductory unit in information systems found that computer anxiety did affect test performance. The students who failed the test reported in this study had substantially higher levels of computer anxiety. Unfortunately it can be argued that in this case computer anxiety is relevant to the content of the test as well as the mode in which it was administered. In his survey of the literature, McDonald (2002) noted that: "it is unclear what effects computer anxiety has on computer-based tasks" (p. 306), although he does go on to note that the "available evidence suggests that computer anxiety can have a negative effect on test scores" (p. 307). Again, this is an area where the relationships are by no means clear-cut.

The notion that increased computer anxiety always produces decreased performance is certainly too simple: the relationship is most likely curvilinear rather than linear (Sapp, 1993); a parabolic, or inverted U, relationship may be more appropriate, with high and low anxiety both producing a lower performance than a moderate level of anxiety. In support of this sort of relationship, students sitting a computer-based form of the Graduate Records Examination who scored most highly relative to a paper-and-pencil version of the examination exhibited a moderate level of computer anxiety (Vogel, 1994). How much anxiety should be regarded as 'moderate' is, of course, the key question, and the answer depends on the difficulty of the test. With a difficult test, a lower level of anxiety is more optimal. An investigation into the effects of item difficulty and test anxiety on the test performance of 90 undergraduate students was carried out using multiple-choice verbal aptitude items (Rocklin & Thompson, 1985). The results showed that: "The least anxious students in the sample did best on a very hard test, and moderately anxious students did best on an easy test, whereas the most anxious students did poorly on both tests" (p 368). Immediate feedback improved performance, especially for students who had done an easy test.

From the above, it can be seen why it is particularly difficult to determine a relationship between anxiety and performance. Consider a group of students who are sitting a test. Some of the students will be very capable and well prepared and will find the test easy. For these students, a moderately high level of anxiety will produce the best performance on that test. However, in the group of students there will also be some students who are ill-prepared to take the test and who consequently will find it very difficult. For these students, a lower level of anxiety will produce the best test performance. The optimal level of anxiety for a test candidate will depend on the candidate as well as the test difficulty.

Another factor to consider in the establishment of a simple connection between computer anxiety and performance on CBA is the nature of the task itself. For example, it is known that anxiety reduces the capacity of working memory (Eysenck, 1988), so it should follow that tasks that require little working memory may be less affected than tasks requiring much working memory. As a generalisation, it may be that more complex tasks are affected by anxiety more than simpler tasks as these can be readily completed without great demands on working memory. However, there appears to be little analysis in the literature following this avenue of enquiry.

Problems in defining and measuring computer anxiety may also contribute to a lack of clarity in establishing a relationship between computer anxiety and performance on CBA. The somewhat disconcerting conclusion drawn by Chua, Chen, and Wong (1999) that "instruments measuring computer anxiety are generally reliable, but not compatible with one another" (p. 609) lends weight to the notion that there may indeed be some problems with definitions of and measuring tools for computer anxiety. Further, an investigation conducted by Tseng, Tiplady, Macleod, and Wright (1998) into how computer anxiety affects assessment performance when that assessment involves pen and paper, a PDA, or a computer produced the strange result that computer anxiety negatively affected both the CBA and the PPA, but not the assessment taken using a PDA. As the authors concluded: "It is not clear why such a relationship should be found, and this observation requires further investigation" (p 608).

There is also debate about whether computer anxiety is a separate entity or really a form of test anxiety. The extent of the uncertainty in this area is illustrated well, and perhaps unwittingly, by Shermis and Lombard (1998) who advanced the belief that what is test anxiety is often mistaken for computer anxiety: "The results suggest that much of what is considered computer anxiety may in fact be a manifestation of test anxiety" (p. 112), but then went on to state that they had found: "no statistically significant relationship between self-reported measures of computer anxiety and test anxiety" (p. 120) and subsequently went on to conclude that "it appears as if the two measures are tapping two different constructs" (p. 120). Certainly, the fact that some commentators have found computer anxiety to adversely affect the results of tests conducted with pen and paper suggests that the area needs further study (McDonald, 2002; Tseng, Tiplady, Macleod, & Wright, 1998).

4.5 The equivalence of paper-and-pencil and computer-based assessment

As learning institutions replace PPA with CBA, it is important to establish the extent to which these two forms of assessment are equivalent. If these two forms of assessment are to be regarded as equivalent then perhaps the first requirement is that changing from PPA to CBA does not affect the grades that students are awarded. Scaling can take care of some discrepancies in the actual difficulty of the test; what is more important is that rank order of the students is not affected. As with many issues in the recent and rapidly changing area of CBA, the picture is not at all clear. For example, in the area of arithmetic reasoning tests it has been found that computers both enhanced test performance (Johnson & Mihal, 1973) and impeded test performance (Llabre, Clements, Fitzhugh, & Lancelotta, 1987). McDonald's (2002) survey of the literature summarises the situation well: "whereas some authors have found sufficient evidence to conclude that CBA and [PPA] can be considered equivalent [2 references given]… more have concluded that this is not the case [2 references given]… or found mixed results [5 references given]" (p. 302).

Mead and Drasgow (1993) conducted a large meta-analysis of the effects of the medium of test administration to compare CBA with PPA. They analysed data from 36 speeded tests. Cross-mode correlation was estimated to be 0.72 for such tests,

from which they concluded that speeded tests did appear to be affected by the mode of administration. Such an effect is understandable: there are a number of reasons why speeded CBA and PPA might not be comparable. The motor activities and the stimulus presentations and response procedures involved are somewhat different: reading what is written on a piece of paper and then shading in a bubble with a pencil clearly differs from reading what is written on a computer monitor and then pressing a key. For example, it has been noted that reading from a screen is slower than reading from a piece of paper (Mayes, Sims, & Koonce, 2001). If differences such as these could be minimised, by techniques such as writing on the monitor screen, for example, then perhaps the mode of test administration would have little effect.

In order to test the above notion, 10 timed clerical tests were given to 411 undergraduate students (Neuman & Baydoun, 1998). An analysis of the results found no differences across modes: "When speeded CBAs follow the same administration and response procedures as [PPA], differences across modes can be minimized" (Neuman & Baydoun, 1998, p. 71).

The meta-analysis mentioned above (Mead & Drasgow, 1993) also included an analysis of 123 timed power tests, and the cross-mode correlation was estimated to be 0.97 for such tests. Therefore, in this substantial sample of studies, the medium of administration was found to have virtually no effect for power tests. From this it was concluded that CBA and PPA can assess the same construct, provided the assessments are well constructed, contain items of varying difficulty, and allow the candidates sufficient time. Mead and Drasgow also noted that: "No difference in equivalence was observed between adaptively and conventionally administered computerized tests" (p. 449).

In situations where computer skills are relevant to test performance, it is to be expected that CBA may not be equivalent to PPA. For example, one small study examined the effect that mode of administration had on school students' performances on multiple-choice and written test questions (Russell & Haney, 1997). In the case of multiple-choice testing, it was found that the mode of administration had little effect. However, for tests involving writing, for students accustomed to writing on a computer, responses written on a computer led to substantially higher

scores than those written by hand (effect size of 0.9). When these students could take the test on a computer, their pass rate was 67%. When forced to take the 'equivalent' test with paper and pen, their pass rate plummeted to 30%. It was concluded that open-ended test items administered on paper may disadvantage students accustomed to writing on computers. Other studies have drawn similar conclusions (Russell, Goldberg, & O'Connor, 2003; Russell & Plati, 2002). Similarly, a study considering the effect of keyboarding speed found that: "for students whose keyboarding speed was at least 0.5 or one-half of a standard deviation above the mean, performing the language arts test on computer had a moderate positive effect. Conversely, for students whose keyboarding speed was 0.5 standard deviations below the mean, performing the tests on computer had a substantial negative effect" (Russell, 1999). In the light of the above, for tests involving written answers, it would seem reasonable to allow students to take either a paper-and-pen or a computer-based version of the test, unless one aim of the test was to assess pen-writing or typing.

Deficient design can make CBA a poor relation of PPA. For example one would hardly be surprised by the result that the marks scored on a paper-and-pencil test were higher than the marks scored on a supposedly equivalent computer-based test when the latter denied students scratchwork space (J. A. Lee & Hopkins, 1985). Similarly, PPA does allow students the opportunity to answer questions in any order, revise and change answers, and it should be a natural requirement of CBA to allow students the same (Wise & Plake, 1989). The American Psychological Association's guidelines for CBA state that it should normally "provide test takers with at least the same degree of feedback and editorial control regarding their responses that they would experience in traditional testing formats" (APA, 1986, p. 12).

OASIS tests and assignments do provide students with the opportunity to review and change their answers. Students are also at liberty to work out the answers to the questions using their normal calculator and pen and paper. Only a small amount of keyboarding is involved. The questions in OASIS tests and assignments cover a range of difficulties from easy to hard. The evidence is that students are able to complete the questions for these assessments well within the time requirement. Therefore, the assessments may be regarded as power assessments rather than speeded assessments. In the light of the evidence above, it would seem reasonable

that OASIS assessments may be regarded as equivalent to the same assessments provided in a paper-and pencil-format.

Of course, there is no absolute requirement that CBA regimes must be completely equivalent to PPA regimes. The assessment programme for a course may well be different if it changes from PPA to CBA: the automation of marking may make more frequent assessment possible. Practice tests may take place. And, just as the overall assessment programme may change, so too may the individual assessments that make up that programme. Most current CBA systems provide far richer data about student performance than do paper-based systems (Conole & Warburton, 2005). A consideration of what data are appropriate and desirable to obtain may well shape the assessment programme for a course. And in turn, the data obtained may well modify subsequent aspects of the course, including assessment items still to come. One extreme example of this is adaptive testing, in which a candidate's performance on test items during an assessment actually determines what test item is presented next. Today's computer technology certainly makes adaptive testing a fully-realisable possibility. Without such technology, it would not be feasible in the current era of large student numbers in tertiary education.

Establishing CBA as equivalent to PPA, and then changing from PPA to CBA should not be seen as the final objective, although, for reasons of credibility and change management, it may be seen as a first step (Raikes & Harding, 2003). CBA provides additional opportunities, and there is no reason why these opportunities should not be taken. At the very least: “direct translation of paper-based assessments into online assessments is inappropriate; there is a need to revisit question formulation, reflecting on what it is intended to test” (Conole & Warburton, 2005, p. 21). CBA can also provide things that PPA cannot (Jacob & Chase, 1992), such as animations and video and sound clips. Computer-based marking may be more reliable than human marking: one study that investigated equivalence of mode of administration for a career decision-making system found that the computer-based version was significantly more reliable than the paper-and-pencil form, and was generally equivalent in other respects (Kapes & Vansickle, 1992).

The easy apparent objectivity offered by CBA may also offer a clear advantage in some situations fraught with emotional or ethical difficulties. For example, there is a pressure on lecturers to maintain a high pass rate, and Hawe (2003), referring to a significant body of literature, noted that, in professional subjects including social work, law, nursing and related health fields, there is much stress associated with the awarding of failing grades and therefore a tendency to award passing grades to poor students in the hope that the next assessor will award them the failing grade that is more appropriate. Further, when lecturers form close working relationships with students, their reluctance to award failing grades is increased (Sabar, 2002). The emotional and subjectivity issues inherent in human marking may be removed by using the automatic marking offered by CBA (Sim, Holifield, & Brown, 2004).

There are also dangers in what is described as objective testing. Such tests may be limited in their choice of formats, relying heavily on multiple-choice. There is some debate about the ability of objective testing to measure higher-order skills. The current view is that objective questions can challenge students, test higher-order thinking skills, and examine a wide range of learning outcomes, at both undergraduate and postgraduate level, provided that the questions are constructed with care and imagination (Bull & McKenna, 2004; Conole & Warburton, 2005). However, objective tests are best used in conjunction with a variety of other assessment methods. There are limits to the skills they can assess: “They cannot, for example, test a student's ability to communicate, to construct arguments or to offer original responses” (Bull & McKenna, 2004, p. 21).

The clear marking schemes generally used in objective tests are in fact two-edged swords. Any mistakes made in implementing marking schemes or in subsequent data management may be readily noticed by students keen to gain extra marks. While this is a good thing, it places extra pressure on academic and administrative staff to execute all operations in any error-free fashion (Conole & Warburton, 2005).

More generally, there are a number of risks involved in implementing CBA. Perhaps the most worrying risk is that of the system crashing during an assessment. Such a crash could result from software problems, server problems, computer terminal problems, or network problems. Since much of the software involved in CBA is

relatively new and therefore relatively untried, and often home-grown, there is a reasonable probability that it could produce problems. For example, student answers may be marked incorrectly, or student scores totaled incorrectly. There are also the normal financial risks associated with implementing a change that may or may not be received successfully. Zakrzewski and Steven (2000) describe CBA as "a 'risky' activity" (p. 202). However, the most severe risks in their view concern things such as: a lack of access to computing resources during CBA periods for students who were not involved in the assessment, academic staff opting out of CBA because of the need to completely redesign tests each year, computer workstations being too close to each other in the test environment, and academic and support personnel with inadequate technical skills. Risks such as workstation and server collapse during examinations were not likely occurrences at the University of Luton, as described earlier in Section 4.2.1.

For CBA to replace PPA, the former must be acceptable to students as well as instructors. There is little point in establishing equivalence between the two if only one is acceptable to students. Fortunately, there is a good level of empirical evidence to suggest that students do find CBA an acceptable if not preferable assessment technique (Chou, 2000; Sambell, Sambell, & Sexton, 1999; Scott & Stone, 1999; Sim, Holifield, & Brown, 2004; Thoennessen, Kashy, Tsai, & Davis, 1997): "We announced that... all assignments could be submitted on paper; most students still chose to use the computer system however" (Scott & Stone, 1998, p. 15). This preference has often translated into increased student hours put into tasks: "The students rated the system extremely favourably even though they spent significantly more time on the assignments compared to traditional classes" (Thoennessen & Harrison, 1996, p. 141). Trust of CBA was also mentioned as a factor: "One major reason for the positive student feedback is the trust of the students in the computer... The students continued to work until they solved the problems because they were convinced that the answers were correct" (Thoennessen & Harrison, 1996, p. 146). Another factor is the prompt feedback possible with CBA. This has already been discussed in Section 3.5.

On this issue of CBA being favoured as an assessment tool, Pascarella (2002) sounds a note of caution. She investigated how CAPA (Computer-Assisted Personalized

Assignments) was used by 500 students in the calculus-based introductory physics course at the University of Colorado, Boulder. In general, students liked CAPA (74%) more than traditional assignments (19.5%). CAPA was also seen as a better learning tool too (57.6% versus 32.8%). However, it appeared to be the non-expert learners who favoured CAPA, while the expert learners favoured traditional assignments. It was simply that most of the learners were non-experts (approx 60%). The expert learners saw traditional assignments as a better learning tool (75% to 25%) and also liked them better (54% to 42%). For non-experts, CAPA was seen as a better learning tool (75% to 11%) and also liked better (89% to 3%). CAPA may have appealed because it offered instant feedback, as opposed to feedback delayed by one-week; and also because it allowed students 15 attempts without penalty on each question, making guesswork, particularly for items consisting of a small set of questions with T/F answers, a viable option. However, the expert learners may have reacted against the 15 attempts rule, regarding it as enabling students with poor skills to score highly. They may also have liked the way that partial credit could be gained for working in traditional assignments. It is interesting to note, from the statistics given above, that there was a significant number of students who preferred CAPA yet saw the traditional assignments as a better learning tool.

4.6 Summary

Three successful implementations of CBA have been discussed in this chapter, each being well received by both instructors and students. Once practices were established, instructors saved time, the amount of student work marked increased radically, and both staff and students benefited from prompt and detailed feedback. Where formative assessment was available, students devoted significantly increased time to their studies and lifted their achievement. The above matches our experiences in developing OASIS.

The issue of computer anxiety is complex because it is not simply related to either computer experience or test performance. Computer anxiety depends on student perception of computer experience, rather than actual experience. Test performance is a curvilinear, not linear, function of computer anxiety, and the parameters of the function depend on the test difficulty, a quantity that varies from student to student. Finally, computer anxiety is sometimes confused with test anxiety.

Computer anxiety would appear to be a minor issue for OASIS: the students who use it are very familiar with computers and are able to use OASIS for practice prior to taking assessments. Furthermore, the assessments themselves are low stakes, and can easily be completed within the time allocated by the vast majority of students.

The experience of sitting OASIS assessments should largely be equivalent to paper-and-pencil tests. Students can review and revise answers, and answer questions in any order. Little typing is involved, and time allocations are generous. The marking, however, is not equivalent, as each answer is either marked right or wrong. No credit is given for working. However, a question can be split into parts, with credit for each part answer. Consequential marking can be used to prevent early mistakes discrediting subsequent correct calculations.

Chapter 5. Methodology

5.1 Purpose and outline of chapter

This chapter concludes the literature review that has been presented in the previous three chapters and forms a sound basis for the presentation of the research findings that begins in Chapter 6. The action research paradigm and its suitability to the present research programme are outlined in the next two sections. This is followed, in Section 5.4, by an exploration of the claim that the action research methodology should be fully accepted as a genuine research methodology. Issues concerning rigour in action research are addressed in Section 5.5. An attempt is made to say what good action research should look like and what criteria it should meet. The overall research aim, and the research goals and their associated research methodologies are presented in Section 5.6. The specific research methods are described in Section 5.7. This section also gives some indication of how these research methods, individually and collectively, contributed towards achieving validity criteria. Some of the practical issues that arose in the implementation of the action-research cycles are highlighted in Section 5.8. The systems that were put in place to ensure that the research was conducted in an ethical manner are outlined Section 5.9. Finally, the chapter is summarised in Section 5.10.

5.2 Introduction

As mentioned previously in Section 1.6, my dual roles of instructor and researcher attracted me to the action research paradigm. I did not set out, simply as a researcher, with the single aim of establishing whether the use of OASIS enhanced student learning. Had I done so, I could simply have compared the performance of a group of students who used OASIS with the performance of another group who did not use OASIS. Such a comparison could have clearly indicated a quantitative difference. However, such an investigation was deemed unethical since there was already some evidence that students did gain from using OASIS. As an instructor, I had also to include the aim of maximising student learning. For this reason, the responsive nature of action research was deemed most suitable: research findings could rapidly lead to modifications in OASIS and its implementation so as to bring immediate

benefits to the students. These ongoing modifications left little room for the replication of the findings from one class in other classes run at later dates. However, this disadvantage was regarded as a reasonable price to pay. A conscious choice was made in favour of responsiveness over replicability.

5.3 The action research paradigm

The social psychologist Kurt Lewin coined the term 'action research' in the United States in the nineteen-forties (Lewin, 1946). He believed research that produced only books was inadequate (Lewin, 1948), and he saw action research as "research which aimed to promote social action through democratic decision-making and active participation of practitioners in the research process" (Kember & Kelly, 1993, p. 1). Recent definitions include the following:

> Action research is true to label -- it is action and research. (Dick, 2002)
>
> Action research is a process by which change and understanding can be pursued at the one time. (Dick, 1997)
>
> [Action research is] the study of a social situation with a view to improving the quality of action within it. (Elliott, 1991, p. 69)
>
> Educational action research is an enquiry which is carried out in order to understand, to evaluate and then to change, in order to improve some educational practice. (Bassey, 1998, p. 93)
>
> The aims of any action research project or program are to bring about practical improvement, innovation, change or development of social practice, and the practitioners' better understanding of their practices. (Zuber-Skerritt, 1996a, p. 83)

The above definitions make it clear that action research really does involve the practitioner in two activities: acting to bring about change and investigating to increase understanding. These two activities are tied together in a recurrent cycle of activities that is sometimes visualized as a spiral. In fact, Lewin (1946) advanced the spiral model of planning, action, observation and reflection which is generally still followed today.

> Action research... is usually described as cyclic, with action and critical reflection taking place in turn. The reflection is used to review the previous action and plan the next one. (Dick, 1997)

> The project proceeds through a spiral of cycles of planning, acting, observing and reflecting, with each of these activities being systematically and self-critically implemented and interrelated. (Carr & Kemmis, 1986, p. 165)

> [Action research follows] a cyclical process of:
> 1 strategic planning;
> 2 action, i.e. implementing the plan;
> 3 observation, evaluation and self-evaluation;
> 4 critical and self-critical reflection on the results of points 1-3 and making decisions for the next cycle of action research. (Zuber-Skerritt, 1996b, p. 3)

> In each cycle there are three major phases: investigation and analysis, planning and action, and evaluation and reflection. The third phase creates the basis for deciding "what next?", and for moving onto a further cycle in which these three stages are repeated, focusing on a new dimension of the original issue or a new issue that emerged during the first cycle. (Cardno, 2003, p. 14)

The literature on action research is in general agreement concerning the defining characteristics outlined above. However, a number of commentators note that the term 'action research' covers a myriad of somewhat different activities (Cardno, 2003; Henry & McTaggart, 1996; Herr & Anderson, 2005). Just what constitutes action research is not universally agreed upon. For example, there is some disagreement about whether the researchers are required to be insiders. On the one hand, Herr and Anderson (2005) see no such requirement for action research, and they use the term practitioner research to denote action research in which the researcher is an insider:

> [The term practitioner research] implies that insiders to the setting are the researchers, whereas in other traditions of action research, the researcher is an outsider who collaborates to varying degrees with insider practitioners or community members. The term action research leaves the positionality (insider or outsider) of the researcher open. The term practitioner researcher places the insider/practitioner at the center of the research. (p. 3)

However, others see action research as research necessarily undertaken by insiders:

> There is a clear understanding that action research is research undertaken by those in the field: field workers, teachers, administrators or supervisors in order to change and improve their own practice. (Kember, D., & Kelly, M., 1993, p. 2)

> Action research is a form of collective self-reflexive inquiry undertaken by participants in social situations in order to improve the rationality and justice of their own social or educational practices, as well as the understanding of these practices and the situations in which these practices are carried out. (Kemmis & McTaggart, 1988, p. 5)

In what follows, the viewpoint that the primary researchers in action research are insiders has been adopted.

There is also some disagreement in the literature about whether action research must be a collaborative exercise. Some commentators do not explicitly raise this issue, while others do but leave the issue open: "It is commonly done by a group of people, though sometimes individuals use it to improve their practice" (Dick, 1997), "it can be undertaken by the individual teacher, a group of teachers working co-operatively within one school, or a teacher or teachers working alongside a researcher or researchers in a sustained relationship" (Cohen, Manion, & Morrison, 2000, p. 226). Herr and Anderson (2005) speak in favour of collaboration: [a practitioner researcher at the center of the research] "often tends to decenter other important stakeholders, such as clients and other committee members. Because of this, many argue that action research should always be collaborative" (p. 3). Kemmis and McTaggart (1988) are quite definite on the matter: "The approach is only action research when it is collaborative" (p. 5).

The above argument by Herr and Anderson highlights one of the potential strengths of action research: it can empower those involved. In traditional research, the researchers are distant from the research setting. Ideally, they make observations without disturbing the system being observed. Any disturbance of the system would traditionally be regarded as making the research less valid. The subjects of the research have the research done to them but derive no direct benefit from it. By contrast, the theme that action research empowers the participant researchers, and improves both their understanding and the situation in which they find themselves, is prevalent in the literature (Cohen, Manion, & Morrison, 2000; McNiff, Lomax, & Whitehead, 2003; Zuber-Skerritt, 1996a).

The above theme of empowerment fits in well with the goals of the present research, which rate most highly the aims of improving the situations in which both instructors and students find themselves. Clearly, these aims would be best achieved through strong collaborations with the instructors and students involved in using OASIS. However, here a note of caution is sounded by Herr and Anderson (2005): "the issue of collaboration and participation creates important tensions in the case of action

research dissertations, because the culture of dissertations discourages collaborative work" (p. 4).

Certainly, in conducting my research, and in line with the collaborative spirit and philosophy of action research, I have entered into collaborative relationships with my colleagues and our students. All those involved with OASIS have been regularly asked for feedback, and surveys and interviews have also been conducted and analysed. Proposals for change have come from a variety of sources, not just from myself, and these proposals have been discussed with students and staff prior to implementation. Beyond this, further collaborations have been entered into in order to obtain specialist knowledge in some areas such as programming and survey design, and also to obtain funding. A committee, with myself as convenor, has been set up by the Head of Department to oversee the implementation of OASIS. However, while this committee may well act on suggestions arising from research findings, I am solely responsible for the conception, direction and implementation of the research programme. As one small example, I have arranged, conducted and analysed all the interviews with students. Of course, without OASIS being implemented across a range of course content areas, and being used frequently by a large number of students, the research programme itself would not have been viable. Since embarking on this research programme, I have designed and written sufficient OASIS questions to cover all areas of all courses that I teach, thus establishing an adequate research base, both in terms of student and question numbers and range. While my work has necessarily involved some collaboration as outlined above, my role in the research programme has been dominant and sufficient on which to base this dissertation.

The theme of empowerment in action research is often taken further and given a political context. McNiff, Lomax, and Whitehead (2003) make the point that action research is intentionally political:

> Action researchers need to understand that they are frequently in potentially politically contested scenarios. When practitioners begin to question the current and historical contexts of a situation, and perhaps begin to reveal injustices, they have to make decisions about whether they wish to follow their own value commitments and try to improve the situation according to what they believe in, or whether they will go along with the status quo. (p. 15)

Grundy (1987) goes still further, seeing action research as entailing a struggle against illegitimate repression, domination and control, and containing the political agenda of moving towards a more just and egalitarian society. This rather extreme view is perhaps more understandable when it is remembered that action research grew out of Lewin's (1946, 1948) work with disadvantaged and marginalised groups. In fact, notions of emancipatory action research similar to Grundy's are often voiced (Carr & Kemmis, 1986; Cohen, Manion, & Morrison, 2000; McTaggart, 1991; Zuber-Skerritt, 1996a).

While Grundy's concept of action research is a step too far for the present research, the point made by McNiff, Lomax, and Whitehead above about the political nature of action research does highlight a real issue. Action research involves change, and this change may be resisted by inhabitants of the research setting or by those who have some power over the research setting. For example, one instructor on a course may wish to change something about the way the course is taught while other instructors may be resistant to change. Other members of the Department, or perhaps the Head of Department, may oppose changes in one course because of the way these changes may affect other courses, or the perceptions of those involved in other courses. Proposed changes may have budgetary consequences, and, given the likelihood that demands on the budget exceed the actual funds available, some political negotiations may need to be entered into in order to successfully effect the changes. If the resistance to change encountered is too great, the proposed action research programme simply cannot proceed. This is not a problem encountered by traditional researchers, particularly those who are outsiders to the research setting, since there is no direct requirement on their research to implement or even suggest change. Furthermore, as outsiders, whether any changes suggested by the research findings are implemented is of relatively little consequence. By contrast, for participants in the research setting, a failure to implement changes, particularly those that their own well-founded research advances as improvements to the setting, would be a severe setback. Fortunately, in the implementation of this research programme, no untoward resistance to change has been encountered, from either my colleagues or the Head of Department. In fact, once I had explained and justified my proposed course of action, I found myself enjoying a good level of support for my activities. Beyond the Department, at university level, I have so far been unable to secure funding to

implement OASIS in high-school physics classes or as a deliverer of entrance tests to prospective overseas electrical engineering students. While both of these projects are desirable, the fact that they have not, as yet, moved beyond the conceptual stage is of no great consequence to the research programme currently being pursued within the Department. Adequate funding for that research programme has been secured.

5.4 Action research as genuine research

Herr and Anderson (2005) describe action research as a "complex and messy process" (p. 7). Dick (1993) claims that action research: "is harder to do than conventional research. You take on responsibilities for change as well as for research. In addition, as with other field research, it involves you in more work to set it up, and you don't get any credit for that". However, while traditional research can improve one's knowledge, action research can, in addition, improve the research setting itself, enhancing the understanding and teaching of the academics and the learning of the students. These attractions may be enough to tilt the balance in favour of action research for a researcher who is also an instructor situated within the research setting. There are certainly arguments both for and against action research. For example, some perceive a problem concerning the rigour of action research. This arises from the fact that action research runs counter to conventional research paradigms, in that it entails changing the situation in which the research is conducted and frequently relies heavily on qualitative data.

Some advance arguments that action research is not really research at all. Such arguments should start with a clear definition of research. Cardno (2003) suggests that "research simply means systematic investigation into a subject in order to establish facts and reach new conclusions" (p. 1). After considering a range of different types of research, Blaxter, Hughes, and Tight (1996) suggest that "the basic characteristics shared by all of these... is that they are, or aim to be, planned, cautious, systematic and reliable ways of finding out or deepening understanding" (p. 5). Bassey (1998) defines research as "systematic, critical and self-critical inquiry which aims to contribute to the advancement of knowledge and wisdom" (p. 38), and then goes on to propose that educational research "is critical inquiry aimed at informing educational judgments and decisions in order to improve educational

action" (p. 39). There would appear to be nothing in any of the above which precludes well thought-out and carefully conducted action research from being regarded as genuine research. In fact, if one aim of educational research is "to improve educational action" as suggested above, then action research is to be favoured over traditional research. Cohen et al. (2000) suggest that "action research is well placed to overcome the perceived persistent failure of research to impact on, or improve, practice" (p. 227).

Perhaps, as Hodgkinson (1957) suggests, action research is second-rate research, a weak form of educational research suitable for busy teachers. He saw it as being conducted by people who lacked adequate training in research techniques and whose methods were sloppy and unsystematic. Hodgkinson saw action research as entailing a commonsense, problem-solving, approach rather than employing scientific method. A genuinely scientific method would involve precise definition, measurement, and control of the variables. In Hodgkinson's attack, there seem to be two issues. One is that the people doing action research are not qualified to do real research. The second is that action research itself is not scientific enough to be real research. The first issue seems to have been taken care of by time. Cohen et al. (2000) note that action research's "combination of action and research has contributed to its attraction to researchers, teachers and the academic and educational community alike, demolishing Hodgkinson's (1957) corrosive criticism of action research as easy hobby games for little engineers!" (p. 226).

But perhaps action research is inherently weak research, no matter who undertakes it. Teachers on a day-to-day basis reflect on their teaching and make changes and improvements as a result of that reflection. Certainly, this activity would not usually be classified as research. Perhaps action research is too similar to this day-to-day activity to be classified as genuine research. Proponents of action research would argue that there are key differences: action research is more systematic and reflective in its approach, for example:

> When action research is explained to people they often respond "Oh, I do that already". And to some extent they do. It is natural to act and then review the results of the action. Much of the time, however, many of us do not reflect as regularly or as critically or as systematically as we might. With regular, critical and systematic reflection we can have more confidence in our research

> conclusions. Without it we may overlook some important evidence. (Dick, 1997)
>
> To do action research is to plan, act, observe and reflect more carefully, more systematically, and more rigorously than one usually does in everyday life. (Kemmis & McTaggart, 1988, p. 10)
>
> It is not the usual thinking teachers do when they think about their teaching. Action research is more systematic and collaborative in collecting evidence on which to base rigorous group reflection. (Kemmis & McTaggart, 1992, p. 21)

It is reasonable that a teacher who is sufficiently systematic and rigorous in reflecting on her or his activity be described as a researcher, just as a person who gazes at the night sky on a regular basis could be described as an astronomical researcher, provided her or his activity was sufficiently systematic and rigorous. In both cases it would be hoped that new knowledge would emerge and bring with it some benefit.

Hodgkinson's (1957) claim that action research is not sufficiently scientific, or quantitative, to qualify as genuine research has also been discredited through the passage of time. These days, provided they are applied correctly in appropriate settings, examples of both quantitative and qualitative research can qualify as genuine research. Further, the fact that quantitative and qualitative research are sometimes known respectively as positivist and post-positivist research suggests that the latter has gained purchase following the identification of some shortcomings inherent in the former (Ary, Jacobs, & Razavieh, 2002; Cohen, Manion, & Morrison, 2000; Gall, Borg, & Gall, 1996). Ary, Jacobs and Razavieh note these problems with applying the scientific approach in education research: the complexity of the subject matter, difficulties in observation, difficulties in replication, interaction of observers and subjects, difficulties in control, and problems of measurement. Nor are all the problems in the application. Cohen et al. note that there is a fundamental difficulty at the philosophical root of positivism as applied to social sciences:

> The difficulty in which positivism finds itself is that it regards human behaviour as passive, essentially determined and controlled, thereby ignoring intention, individualism and freedom... The more effort, it seems, that researchers put into their scientific experimentation in the laboratory by restricting, simplifying and controlling variables, the more likely they are to end up with a pruned synthetic version of the whole, a constructed play of puppets in a restricted environment. (p. 19)

Wragg (1994) offers us a pithy reality check: “While the counting of events may offer some interesting insights it falls far short of telling the whole story of classroom life” (p. 9).

In fact, as is now well accepted, there are difficulties with the traditional notion of scientific discovery in general. First, there are philosophical difficulties. For example, the notion of keeping an open mind is incompatible with deciding what variables are important and formulating a hypothesis to test. Second, the traditional vision of scientific knowledge as growing carefully and rationally through ongoing data collection and more accurate experimentation carried out by dispassionate scientists does not match what actually happens in science (Kuhn, 1970).

It is not that one research paradigm is good while the other is bad; rather they entail different philosophical bases and methods. Which is preferable depends on the situation. Further, contrary to Hodgkinson’s criticism, action research is not limited to qualitative methods. That action research frequently does involve qualitative methods is not a sign of weakness, but rather an indication that action research is frequently applied to settings in which it is particularly appropriate to choose such methods. Further, quantitative and qualitative research are to some extent complementary, and there is much to be gained by employing both research methods. The action research programme described in this dissertation does rely on both quantitative and qualitative methods. As a research programme, it will stand or fall on the merits of these quantitative and qualitative research methods, not on the fact that it is a piece of action research. As Cohen et al. (2000) remark in their survey of research methods in education, action research is “a flexible, situationally responsive methodology that offers rigour, authenticity and voice” (p. 241).

One argument that may be levelled against action research concerns the conflicts of purpose and activity between research and teaching. Research and teaching have different primary purposes; for research it is to learn through investigation, for teaching it is to bring others to understand. Since these primary purposes are different, research and teaching entail different activities. Therefore, some conflict may be expected when one individual takes on both research and teaching roles. Further, if the teaching role cannot be set aside when such conflict does occur, the

quality of the research is likely to suffer. Wong (1995) gives an example of a conflict that occurred when he was both teaching and researching in a science classroom. The aim of the research was to determine the extent to which the students' explanations of natural phenomena were 'scientific'. A conflict arose when, as a researcher, he wanted to question in some detail a student about her explanation of an event while, as a teacher, he realised that his extended questioning was making the student feel uncomfortable in front of her peers who were in turn losing interest in the lesson because they felt that they were not receiving enough attention. Wong suggests that this sort of conflict could be reduced if a classroom culture were established in which "extended teacher-student interactions were perceived as a normal, valuable part of the classroom experience" (Wong, 1995, p. 26). Such a culture would close the gap between research and teaching activities. However, he notes that establishing such a culture would not be easy. Wong's example highlights a problem faced by many teacher-researchers.

However, even assuming no compromise in teaching and learning outcomes, the conflict of purpose and activity between research and teaching need not lead to poor research outcomes. This is because there is no absolute requirement that the research and teaching activities take place at the same time. In Wong's case, perhaps he could have questioned the student in more detail when the rest of the class was involved in experimental work or problem-solving, or even when the class had finished. In the case of my research, the teaching and the research do happen largely at different times (to be fair, this is easier to achieve in a university setting as, in a school setting, much of the day is taken up with direct student contact). Apart from about 10 minutes per semester devoted to completing a survey form, my lectures, laboratory sessions and tutorials are devoted to teaching. Much of the research involves analysis of data, an activity that does not directly involve students. The bulk of the research that does directly involve students takes place during interviews. While my teaching and research activities are different, they take place at different times, and one does not reduce the effectiveness of the other. As a teacher and also as a researcher, I aim to improve student learning. Teaching and learning is the foundation on which the research is based. The fact that the teaching and the research are both done by the same person does not reduce the quality of either but, rather, improves both.

As an education research methodology, action research offers three important attractions to instructors. First, it enables them to develop as reflective practitioners, gaining knowledge about their own practice and deriving and developing their own educational theories from that practice. It has been argued strongly that systematic reflective practice is an effective way for instructors to learn (Schon, 1991). Second, it places teachers at the forefront of the research rather than relegating them to a minor role of research object. It makes perfect sense that education research should recognise the important place that practitioners hold in the research setting. They are immersed in the setting, they are knowledgeable about the setting, and they will value highly findings from research in which they were meaningfully involved. Frequently the gap between researchers and practitioners is too wide, and the findings of the former group are ignored by the latter group (Costello, 2003; Rose, 2002). Third, action research enables practitioners to implement changes in their circumstances, changes they have faith in and a stake in since they were justified by research in which the practitioners themselves had a key role. Through implementing such changes, instructors can improve the quality of teaching and learning and also the conditions of work in their institutions. The prospect of such improvements is a very strong motive for practising action research (Altrichter, Posch, & Somekh, 1993). These three attractions have no doubt helped increase the popularity of action research, and persuaded academics such as myself to adopt it as a research methodology.

5.5 Rigour in action research

It has been argued that action research is an application of the scientific method to the specific setting of organisations with the specific purpose of organisational change (Arguinis, 1993). It has also been suggested that action research may be thought of as a complement to and extension of scientific method (Whyte, Greenwood, & Lazes, 1991). In spite of these attempts to close the gap between action research and scientific method, the notion of rigour traditionally associated, at least in principle, with scientific research cannot be sensibly applied without modification to action research. There are two obvious reasons for this. First, action research involves change, so the notion of repeating experimental trials in the same setting becomes highly problematical. Second, action research is typically insider

research, so there is a strong link between the researcher and what is being researched. At the very least, there are major difficulties with the notion that the researcher will not have any effect on the research setting. The researcher will be an important part of the research setting, and the researcher may actually be investigating him or herself. However, while action research does not fit the mould of traditional science, it does provide a valid way to conduct research and it does produce knowledge, a different sort of knowledge to that commonly understood to be generated by traditional science. Its merits cannot be judged solely by the positivist paradigm and a new epistemological basis may even be needed.

> Action research appears as an alternative epistemology based on radically different ontological premises than those of the positivist view on social science, rather than being just another research method... Researchers cannot use positivist methods for assessing the scientific merits of action research. (de Cock, 2005, p. 374)

> But, of course, for any problem of interest to teaching and learning, insofar as it arises and is studied in the actual contexts of practice, one cannot establish true control groups, create random assignments, eliminate potentially confounding phenomena, or, in general, meet the standards of normal-science rigor. Hence, there can be no such thing as a "scholarship of teaching" unless we can change the rules that govern what counts both as legitimate knowledge and as appropriately rigorous research into teaching and learning... The new forms of scholarship call for a new institutional epistemology. (Schon, 1995, p. 34)

> Action research constitutes a kind of science with a different epistemology that produces a different kind of knowledge, a knowledge which is contingent on the particular situation, and which develops the capacity of members of the organisation to solve their own problems... As a procedure for generating knowledge, we believe it has far greater potential than positivist science for understanding and managing the affairs of organisations. (Susman & Evered, 1978, p. 601)

Given the above arguments that action research and traditional scientific research do not share the same epistemological foundation, the following question must now arise: 'How is rigour to be achieved in action research and how do we tell that it has been achieved?'

Dick (1993) suggests the following actions help achieve rigour:

- using a cyclic approach, with each cycle involving data collection, interpretation, and literature search;
- as far as possible working at any time with two or more sources of information ("dialectic"); and

- testing... interpretations stringently by searching out exceptions to the explanations, and explanations of the ambiguities.

He also notes that brief cycles should be used in order to provide adequate iteration. In the present case, each cycle was one semester long. Consequently it was possible to have eight iterations of the research programme.

Robson (2002), in a similar vein to Dick above, suggests the following strategies for dealing with threats to the validity of a piece of research: prolonged involvement in the study, triangulation, negative case analysis, and the maintenance of an audit trail. Prolonged involvement may help reduce respondents' bias as, over time, a more trusting relationship can develop between researcher and respondents. Also, if the researcher is in the research environment for a considerable time, the "novelty value" may wear off and the respondents may behave more naturally. Against this, it must be noted that, if the researcher is also a lecturer and the respondents are his or her students, then there may always be some sort of bias in operation. However, this bias is unlikely to be entirely in one direction. For example, some students may paint a more positive picture than is warranted because they wish to please the lecturer, while other students who are negatively disposed towards the lecturer or the course may paint a negative picture. Triangulation and repetition of the research cycle may help reduce the chance that such bias leads to unwarranted conclusions. Costello (2003) also speaks in favour of all four points advanced by Robson, and notes that "In addition, an audit trail offers evidence that you are being careful, systematic and scrupulous about your research. These are important considerations when you are making the case for its reliability" (p. 46).

Cardno (2003) suggests the following procedures to establish validity:

- Providing accounts of the validity processes in action research reports
- Negotiating the agenda for research and action with all participants
- Using triangulation (of methods, perceptions, observations and interpretations)
- Using cycles of feedback to validate results, tentative analyses, interpretations and conclusions
- Establishing clear audit trails to demonstrate the iterative validation of data by participants
- Implementing iterative processes for dialogue to critique literature, data and change action and the action research methodology itself. (p. 55-6)

In Cardno's suggestions above there are a number of procedures previously mentioned, such as repeated cycles of action, triangulation, and documentation and clear audit trails. It is also made clear that data, interpretations and conclusions should be appraised critically. Additionally, the theme of action research as a collaborative enterprise may be noted in the references to cycles of dialogue and feedback. The suggestion is that validity may be increased through regular and ongoing feedback of results and conclusions to participants. The purpose of such feedback would not be to present them with definitive conclusions but rather to check that the conclusions were in fact trustworthy. In the present research programme this was done in lectures, interviews, focus-group discussions and some informal discussions with students, and also in interviews and many informal discussions with the instructors using OASIS.

Much of the above relates to knowledge generation and research methodology. However, action research is also about changing the research setting and empowering individuals in that setting. Therefore, further criteria are needed in judging the quality of action research. Herr and Anderson (2005) list five goals of action research and identify validity/quality criteria for each. Table 5.5 (based on Herr and Anderson's Table 4.1, p. 55) sets out these goals and criteria.

Table 5.5
Herr and Anderson's goals of action research and validity/quality criteria

	Goals of action research	Quality/validity criteria
1.	The generation of new knowledge	Dialogic validity Process validity
2.	The achievement of action-oriented outcomes	Outcome validity
3.	The education of both researcher and participants	Catalytic validity
4.	Results that are relevant to the local setting	Democratic validity
5.	A sound and appropriate research methodology	Process validity

It should be noted that the above criteria were proposed by Herr and Anderson in a tentative fashion, in part to catalyse dialogue amongst academics and practitioners about what does constitute validity in action research. To the extent that it is

appropriate, these criteria will be applied to the present research. Each of these criteria will be now outlined and discussed in turn.

Outcome validity concerns the extent to which actions occur that solve problems or improve matters in the research setting. Here there is some overlap with democratic validity for the reason that judgments as to whether matters have been improved in the research setting should be made in collaboration with those involved. This criterion of validity is called 'workability' by Greenwood and Levin (1998). Brooks and Watkins (1994) suggest that 'skilfulness' for action research is the equivalent of 'validity' for positivist research and 'credibility' for naturalistic enquiry.

Process validity is defined by Herr and Anderson in a very broad sense. Here we find much of what has been covered previously in this section; for example, the use of triangulation and a series of reflective cycles in order to produce quality evidence. However, process validity also looks at the extent to which the research programme allows ongoing learning of those involved and promotes quality relationships amongst participants.

Democratic validity, as described by Herr and Anderson,

> refers to the extent to which research is done in collaboration with all parties who have a stake in the problem under investigation. If not done collaboratively, how are multiple perspectives and material interests taken into account in the study... While process validity depends on the inclusion of multiple voices for triangulation, democratic validity views it as an ethical and social justice issue. (p. 56)

An action research project that meets the requirement of democratic validity should produce change and knowledge that relates to and benefits the participants. Watkins (1991) speaks of 'relevancy' or 'applicability' criteria. In a similar vein the terms 'local validity' (Cunningham, 1983) and 'ecological validity' (Tandon, Kelly, & Mock, 2001) have been used for criteria that judge how relevant the research outcomes (in terms of both knowledge and action) are to the participants.

Catalytic validity, when "put neutrally... simply strives to ensure that research leads to action" (Cohen et al., 2000, p. 111). However, it is not really a neutral concept, but one that comes with an agenda: to help participants to understand reality in order to

transform it, or to reaffirm their support of the status quo (Lather, 1986). Catalytic validity is similar to some of the types of validity mentioned previously but adds a new dimension: "while this criteria overlaps with process and democratic validity, it highlights the transformative potential of action research" (Herr & Anderson, 2005, p. 57).

Dialogic validity concerns the quality of research as monitored by peer review. This could take place through publishing the research in academic journals, through debate with like-minded researchers, or even through discussions with a person who is prepared to maintain a critical stance.

As Herr and Anderson (2005) comment, "all of these validity criteria for action research are tentative and in flux" (p. 57). Perhaps, in the future, some level of agreement will be reached concerning which criteria are relevant to given circumstances. Until then, "each inquirer must search for and defend the criteria that best apply to his or her work" (Connelly & Clandinin, 1990).

Of the five validity criteria given above for action research, perhaps the most important two for our purposes are outcome validity and process validity. Saying that the present research programme has outcome validity is, in essence, simply saying that it achieves its goals. For example, if student learning is enhanced and instructor workload is reduced, then the programme is well on the way to achieving outcome validity. Of course, correctly judging whether the goals have been achieved relies on process validity. For example, data should have been collected in a variety of ways and from a number of groups of students who used OASIS. In other words, making credible statements about outcome validity depends on meeting the criterion of process validity.

In this particular research programme, democratic validity is to a large extent covered by outcome validity and process validity. This is because the programme aims to benefit the parties that are involved in it, namely the students and the instructors. Further, in order to determine how to best deliver this benefit and to verify that it has in fact been delivered, the viewpoints of students and instructors must be regularly sought. Perhaps it may be argued that this research programme

cannot achieve full democratic validity because the research is not being conducted in a fully collaborative fashion. As mentioned previously, the extent of collaboration is restricted by the requirement that the research be presented for a doctoral dissertation.

A relatively mild, neutral interpretation of catalytic validity will be used for the present research programme. Here catalytic validity will be interpreted in terms of action and increased participant understanding, rather than reality transformations. With the above interpretation, the notion of catalytic validity reduces to elements of the previous three validity criteria and there is no longer any need to deal with it separately. Similarly, little further attention will be paid specifically to dialogic validity. However, here it should be noted that my colleagues and I are in regular discussions concerning OASIS and its implementation, and that research findings concerning OASIS have been presented to the general public in peer-reviewed journal articles and conference presentations (Hussmann & Smaill, 2003; Smaill, 2004, 2005; Smaill & Hussmann, 2003).

The case will be made in Chapter 10 that the research programme has adequately met the relevant validity criteria as outlined above. The making of this case is deferred because, in order to show that outcome validity has been achieved, the outcomes must first be presented.

5.6 Research goals and associated research methodologies

The overall aim of this research was *to develop, implement and validate a Web-based software package that, through providing practice and assessment opportunities, improves student learning and reduces marking and related mundane aspects of instructor workload.* This overall aim was translated into four research goals. For each goal, research questions were written in order to help focus on the goals and determine the extent to which they had been achieved.

The following four subsections present each of the four research goals in turn, together with their associated research questions. The methodology used to address each research question is also stated. Details of the research methods may be found

in Section 5.7. Owing to the spiral, iterative nature of action research, the research goals were visited during each of the eight cycles of the action research programme. These cycles are the subjects of Chapters six to nine. In Chapter 10, I discuss the extent to which these four goals were achieved.

5.6.1 First research goal

To produce a software package (OASIS) that meets certain criteria (e.g. Web-based, secure, robust, scalable, easy to use, provides prompt feedback, logs student activities and results) and can be used for student practice, assignments and tests in electrical engineering.

The four research questions associated with this goal are presented in Table 5.6.1.

Table 5.6.1
Research questions associated with first research goal

	Research question	Methodology
1a.	What criteria are appropriate for OASIS to meet?	Literature review Discussion with colleagues
1b.	Does OASIS meet the criteria?	Direct observation Informal feedback from users Surveys, interviews
1c.	What student activities and results are appropriate for OASIS to log?	Literature review Discussion with colleagues
1d.	Does OASIS log the required student activities and results?	Direct observation Informal feedback from instructors

The literature reviews mentioned above that helped shape the software design are reported in Chapters two, three and four. The purpose in asking the four research questions was to ensure that OASIS did develop as optimally as possible. It was necessary to address the questions in every cycle of the research programme because OASIS, its implementation and users' perceptions all evolved through time. However, the first research goal was under particular focus at the end of 2002 when it was decided that the existing software did not meet the criteria adequately and a complete re-write was necessary. These events are the subject of Section 6.5.

5.6.2 Second research goal

To incorporate features in OASIS that encourage staff to adopt it for their use with courses they teach. In particular, use of OASIS should reduce staff workload. This goal also includes the longer-term goal of extending the use of OASIS beyond the department, faculty and University.

The four research questions associated with this goal are presented in Table 5.6.2.

Table 5.6.2
Research questions associated with second research goal

	Research question	Methodology
2a.	Which instructors are using the package and in what courses?	Direct observation
2b.	Why did these instructors decide to use OASIS?	Informal feedback from instructors Interviews with instructors
2c.	What features of OASIS are appreciated by instructors and what features need changing, deleting or adding?	Informal feedback from instructors Interviews with instructors
2d.	What effect is the use of OASIS having on instructor workload?	Informal feedback from instructors Interviews with instructors

The four questions above were addressed informally in each research cycle. They were also addressed formally in two sets of interviews with the instructors who were using OASIS. The first set of interviews took place at the conclusion of the third research cycle and is described in Section 7.8. The second set of interviews took place at the conclusion of the eighth research cycle and is described in Section 9.6. Background information concerning these interviews is located in Subsection 5.7.6.

5.6.3 Third research goal

To implement the use of OASIS in such a way that student learning is enhanced.

This goal had five research questions associated with it, as presented in Table 5.6.3.

Table 5.6.3
Research questions associated with third research goal

	Research question	Methodology
3a.	How do students use OASIS?	Direct observation Analysis of OASIS data Surveys and interviews Informal feedback
3b.	What features of OASIS are appreciated by students and what features need changing, deleting or adding?	Surveys and interviews Informal feedback
3c.	What student OASIS usage patterns correlate with increased student learning?	Analysis of OASIS data Analysis of assessment results Surveys and interviews Informal feedback
3d.	How do instructors implement the use of OASIS?	Direct observation Interviews Informal feedback
3e.	What instructor OASIS implementation patterns correlate with increased student learning?	Analysis of OASIS data Analysis of assesment results Interviews Informal feedback

These questions were addressed in each research cycle. Student interviews were used in six of the eight research cycles. These are described in Sections 7.7 and 9.5, and Subsections 8.2.3, 8.3.5, 9.2.2 and 9.3.3, with background information presented in Subsections 5.7.2 and 5.7.3. Student surveys dedicated to OASIS were used in three research cycles. These are described in Subsections 7.4.3, 7.6.3 and 8.2.4, with background information presented in Subsection 5.7.1. Because I was in daily contact with instructors, I judged that two sets of instructor interviews would be sufficient (see Sections 7.8. and 9.6). Informal feedback from students was received on a regular basis via email, conversation and course survey forms.

5.6.4 Fourth research goal

To identify best-practice strategies for the implementation and use of OASIS.

The two research questions associated with this goal are presented in Table 5.6.4.

Table 5.6.4
Research questions associated with fourth research goal

	Research question	Methodology
4a.	How should instructors implement and use OASIS to best promote student learning?	Analysis of previous research findings
4b.	How should students use OASIS to best promote their learning?	Analysis of previous research findings

In posing these questions, it was hoped to identify best-practice strategies for the implementation and use of OASIS through a synthesis of the findings associated with research questions 3c and 3e.

5.7 Research methods

The range of research methods was restricted because the research was conducted in a real learning and assessment environment. Research methods that would produce any significant distortion of this environment had to be discounted on the grounds that they would produce results lacking in credibility, as well as on ethical grounds. Sly (2000), while investigating the effect of on-line practice tests on student learning found herself in a similar situation and acknowledged similar restrictions:

> Although the research questions could be answered more effectively using a hypothesis-testing, experimental design, it was not possible (and for ethical reasons not desirable) to manipulate the testing environment for students. Consequently, the research design takes an eclectic approach, using both quantitative and qualitative methods of data collection, attempting to accumulate sufficient data from several sources to provide answers to the research questions. (p. 56)

My research also took an eclectic approach, using a variety of qualitative and quantitative methods. The diversity of research methods used was judged to be sufficient to achieve good triangulation and, consequently, rigour in the research

findings. The rigour was further strengthened by the fact that the research programme ran over eight university semesters, enabling eight cycles of the action-research programme to take place. Over the course of these cycles, the software package itself underwent considerable modification as did the way in which it was used as a learning and assessment tool. Consequently, there was also some modification of the research programme during this period of time. Furthermore, issues that were identified by students in early interviews were followed up in later interviews and frequently led to modifications in the software or its implementation.

The following sub-sections discuss sources of data and associated research methods.

5.7.1 Surveys

The survey used in this research grew out of a survey that had been used previously to determine how year-two students felt about two software packages: Cecil, described briefly in Section 1.1, and an earlier version of OASIS itself (Hussmann & Bigdeli, 2002; Hussmann & Smaill, 2003). The pre-cursor survey consisted of the following eight statements, to be rated by students on a five-point "Strongly agree" to "Strongly disagree" scale:

1. I often access Cecil and OASIS.
2. I come to campus less often because of Cecil and OASIS.
3. The tools provided are easy to use.
4. I like the instant performance feedback using OASIS.
5. OASIS helped me to prepare for the tests.
6. OASIS helped improve my skill level.
7. I like the supplementary material within Cecil (handout, tutorials, marks etc.).
8. The communication between lecturer and student was improved by using Cecil.

Taking Cecil out of the picture reduced these statements to six in number. Further statements were added, increasing the number of statements to 10 in all, as listed below. These statements were to be rated on the same five-point scale used before.

1. OASIS is easy to use
2. I often use OASIS
3. OASIS helped improve my skill level
4. I am more confident about my learning after using OASIS
5. The problems provided in OASIS helped me understand the course material better
6. OASIS helped my prepare for the assessments
7. I come to campus less often because of OASIS

8. I like the instant performance feedback using OASIS
9. I prefer to do problems from the text rather than OASIS problems
10. It would be a good idea to have OASIS in other courses

A second side was also added to the survey sheet and this contained the following five open-ended questions, designed to give students an opportunity to express their ideas and opinions about OASIS in their own words:

1. What do you like most about OASIS?
2. What do you like least about OASIS?
3. What changes or improvements would you suggest for OASIS?
4. What benefits do you think students get from using OASIS?
5. Have you any other comments to offer about OASIS?

The above 15 items were developed and put together after some ongoing discussion had taken place with an academic from UoA's Centre for Professional Development (CPD), one whose research area was CAL. It was considered that her expert input would help produce a survey that would validly and reliably produce a useful range of data. An example of the survey form used is presented in Appendix one.

Two somewhat different approaches were tried in disseminating the survey forms to students and collecting back the completed forms. The first involved handing out the forms at the end of a lecture. While this approach produced a reasonable return rate it was considered that some students may have devoted only a small amount of time and thought to the survey since they may have been in a hurry to get to some other engagement. The second approach involved handing out the forms at the start of a laboratory interview session. This second approach had the advantage that the students were waiting for their own interview so were not in any particular hurry. It also produced a better return rate than the first approach, virtually 100% in fact. However, the students may have had their minds on the upcoming interview rather than on OASIS.

As suggested above, there are two potential problem areas with surveys. The first is the difficulty of getting a full and considered response from the students. The students may well devote only a small amount of time and thought to completing the survey, if they complete it at all. This problem is aggravated by the fact that Faculty of Engineering students are very frequently surveyed; each student is surveyed in each course at least once. These surveys tend to take place towards the end of the

course, also the preferred time to survey the students about OASIS. In an effort to make the students take the OASIS survey more seriously, I always spoke briefly about the great value of the information produced by the survey just before the survey forms were given to the students. In giving such a speech, I was aware that it could skew the survey by appearing to elicit positive feedback about OASIS. I tried to minimise any such affect by thinking about and choosing my words and style of delivery carefully. I was also aware that responses could be skewed somewhat by the mere fact that I was collecting the survey sheets from the students after they had completed them. In order to avoid this problem, I sought help from graduate students and explained that the completed survey sheets would be collected by them. I then left the room and did not return, making it clear that I could not possibly match students to survey forms. However, whether I collected the completed survey forms did not appear to have a significant effect on the way in which they were completed.

The second problem is the difficulty of obtaining a set of survey forms that is representative. The anonymous nature of survey forms is a two-edged sword. On the one hand, it could be argued that students are unlikely to complete survey forms honestly if the forms are not anonymous. On the other hand, if the survey forms are anonymous, it is very difficult to know whether or not a representative group of students has actually completed them. Of course, if virtually every student completes a survey form, then the group can certainly be regarded as representative. However, in contrast to the high-school situation, it is difficult to achieve such a high return rate in university lectures where hundreds of students may be involved. Furthermore, any obvious coercion is likely to skew survey results significantly! As previously mentioned, I did try handing out the survey forms at the start of a laboratory interview session. The students then completed the survey forms and handed them in to the person conducting the laboratory interview at the start of their interview. This approach, unsurprisingly, produced a very high return rate. However, this approach also produced results that did not differ significantly from those obtained in surveys conducted at the conclusion of lectures, where the return rate was significantly lower.

The analysis of each survey was undertaken by CPD. There were two reasons why I did not perform this analysis myself. First, students answered 10 of the survey questions by shading in the appropriate bubble for each question. Manually

extracting the data from hundreds of sheets for these 10 questions would have been an extremely arduous task. By contrast, CPD possessed a scanner that could carry out the analysis much more quickly. Second, it was considered that it would be more credible to have an independent and impartial individual analyse the data from the five open-ended questions. After CPD had carried out the analysis, the survey sheets were always returned to me so that I could look over every individual open-ended response if more detail was needed or some piece of information needed checking.

5.7.2 Interviews with students

A number of students were formally interviewed on an individual basis about OASIS to ascertain their feelings and experiences. These interviews always took place in my office. Each student was asked at the start of the interview if he or she would consent to the interview being recorded. Such consent was always given. All interviews were recorded using a digital voice recorder (Olympus DS-330). The formal, recorded, part of each interview would typically be of approximately 15 minutes duration. However, after I had stopped recording, the interviewee would sometimes make a particularly insightful or revealing comment. In such cases I would try to record on paper as faithfully as possible the essence of the comment.

After each interview I uploaded the digital file of the interview to my computer. I then played back part of the interview to check the recording and uploading. No problems were ever found in this regard. I had transcribed interviews previously, and had found it extremely time-consuming. Fortunately, research funding facilitated third-party transcriptions. I provided the transcriber with digital files of the interviews on a CD or as an e-mail attachment. After transcription, I received a Word file of each interview via e-mail attachment. I then played my file of the interview to check the transcription and make any corrections. Corrections were needed for every interview, due in part to the use of specialist terminology with which the transcriber could not be expected to be familiar. Phrases like ‘Thevenin equivalent’ and ‘node-voltage analysis’ proved popular stumbling blocks.

Normally I began the recorded part of each interview by asking the student to tell me about her or his experiences with OASIS. In doing this, the aim was to allow the student the maximum freedom with his or her comments. However, prior to the first

interview I did prepare a list of questions so that I had something to fall back on if the student was not very forthcoming and I was suddenly lost for words myself. Occasionally I would refer to the list, either to prompt myself, or to check that I had covered all the areas that I had intended to cover. The questions are listed below:

1. In your studies, you have had OASIS problems and textbook problems. In general, did you prefer one to the other? If so, why?
2. Did you do some OASIS problems several times? If so, why?
3. Did you feel that your grasp of the subject improved as a result of using OASIS? Could you tell me more about that?
4. Was your use of OASIS determined by what you were covering in lectures, or by the next test coming up?
5. Did you ever use OASIS in a group situation, or did you always use it on your own?
6. What OASIS features did you like?
7. What OASIS features are not so good?
8. Are OASIS assignments better, or are regular assignments better? Could you tell me more about that?
9. How would you feel about OASIS tests? [As opposed to assignments]
10. What percentage of a course should be assessed by OASIS?
11. What would be the most effective way for students to use OASIS?

While these questions were used as a framework for interviews, particularly earlier interviews, they certainly did not cover the whole range of student responses. Some interviews departed considerably from any script suggested by the above questions. In general, interviews provided an opportunity to explore in more depth issues raised by students in their survey responses, and to follow up any issues of particular concern to students; for example, what level of student collaboration is acceptable for OASIS assignments.

Once several interviews had been conducted and transcribed, the next step was to analyse the data the transcripts contained. For this purpose, the software package N6 (formerly known as NUD*IST) was used. This software package is designed to assist qualitative research and analysis, and it is promoted (N6, 2006) as being particularly appropriate for action research and literature reviews. In fact, I also used N6 for my literature review and can confirm its great usefulness for such a task. N6 was used to classify student comments made during interviews and facilitate the search for patterns in them. This process was ongoing, and the results of the analyses of earlier interviews naturally affected the direction of later interviews.

Obtaining representative groups of students to interview proved to be highly problematical. For my first set of interviews I had carefully gone through student records to compile a list of students that covered the full range of achievement levels and the full range of OASIS usage levels. I then e-mailed the students on the list explaining that I would like them to take part in a short interview about OASIS. However, the response to this method was below 50% and repeated follow-up e-mails brought little improvement to this situation. The group of students who were actually interviewed tended to be those who were achieving satisfactory or better results academically and who were, at least to some extent, making use of the practice opportunities provided by OASIS. Roughly speaking, no student who was in the lowest quarter for academic achievement or for OASIS usage responded positively to a request for an interview. No solution to this problem was found. It was felt that any coercion beyond repeated e-mail requests for interview participation would be unethical and would, most likely, lead to skewed results.

An alternative to inviting pre-selected students to be interviewed was then tried. An email message was sent to all students in the class inviting them to take part in an interview about OASIS. This approach appeared to be just as successful as the previous approach. Surprisingly, it did not appear to produce a significantly less representative group of students than the previous approach. It also provided students who were keen to discuss some issues with OASIS a ready avenue for doing so. Such students were largely missed by the previous approach. The call for volunteers to be interviewed did not produce a group of students who were all 100% enthusiastic about every aspect of OASIS. In general they did feel positively about it, but not more so than the general student population, as determined by survey results. The volunteer student interviewees certainly at times did have negative things to say about OASIS and the interview situation gave them an opportunity to elaborate fully on these issues. If the volunteer group was less representative than the previous group in any way, it was that they tended to be more outgoing and keen to talk about their experiences.

Of course, it should be noted that the previous group had been somewhat self-selecting anyway, since less than half the students who received a personal invitation to be interviewed actually responded positively. It was felt that e-mailing all students

in the class and asking for volunteers was preferable to personally e-mailing selected students and inviting them to take part in interviews, in the sense that the former method resulted in a group of students who produced better information about OASIS usage. For this reason the former method was adopted. In making this decision, it was acknowledged that it would not be feasible to obtain a representative group of students to be interviewed. Interviews would provide a reasonably broad and in-depth picture of how students felt about and used OASIS, while other sources such as surveys and the information collected by OASIS itself would be used to provide data more representative of the entire population, and to check that the information emerging from interviews remained consistent with the overall picture.

In each class there were some students who were not prepared to take part in an interview but who were prepared to put some of their thoughts about OASIS in an e-mail reply. More recently I have been encouraging such replies from students. The text of a typical e-mail request sent to students asking for interviewees is given below:

> Dear students
> I am keen to hear how you found using OASIS this semester. If you would like to take part in a short interview about OASIS, that would be excellent. Just reply to this email and we will organise a time.
> If you don't wish to be interviewed but could email me some of your thoughts and suggestions that would be much appreciated too. I am keen to hear your thoughts about OASIS.
> Thank you for your time
> Chris Smaill

5.7.3 Focus-group discussions

Much of what has been said about student interviews also applies to focus-group discussions. At the start of each discussion, permission to record the discussion was sought from each student individually. The interviews were recorded, transcribed, checked and analysed in the way previously outlined. As the facilitator of the discussions, I tried to play only a minor role in them, aiming to keep the conversation wide-ranging and free-flowing. As for the individual interviews, I occasionally used a set of written questions as a prompt. Where only a few students were involved in a discussion group, it was possible to hold the discussion in my office around my desk. This arrangement proved to be quite successful and had the advantage that all

speakers had to sit fairly close to one of the two microphones being used. For larger groups, a more spacious venue was arranged, such as the Faculty's Boardroom or the Department's Briefing Room. The maximum group size was eight, so these venues were somewhat bigger than necessary.

It was found that focus-group discussions always ran longer than individual interviews, the maximum duration being almost one hour. This may seem like an excessively long time, but in fact the students were, for the most part, actively and enthusiastically involved in the debate. Fruit juice and biscuits were also provided. At best, these discussions were very much student-centred with students raising issues and debating points while my role was extremely minimal. Issues were raised that had not been raised in individual interviews. The depth and extent of discussion was much greater too. At worst, the discussions tended to be led by myself: I would ask a question and the students would each respond to it briefly in turn with only a minimal amount of student-student interaction. Fortunately, there were no discussions that spent the majority of the time operating in that mode.

Following my experiences in attempting to recruit students for individual interviews, I decided that sending out an e-mail to all students in the relevant classes would be the best way to recruit students for focus-group discussions. As before, I accepted that it would not be feasible to gather a fully representative group of students. Rather, the aim was to explore issues in more depth than would be possible with a class survey. It was considered that it would be less intimidating to refer to 'an informal discussion' in the e-mail sent out to students: the phrase 'focus-group discussion' intentionally did not appear. The text of a typical e-mail is given below:

> Dear students
> I'd like to find out more about what you think of OASIS and how you use it, or why you don't! For this reason I'm calling for volunteers to take part in an informal discussion about OASIS Wednesday 27 April at 10 am. If you would be prepared to share your experiences and ideas that would be excellent and I would greatly appreciate it. Please email me and I'll get back to you confirming location, etc.
> Thank you for your time.
> Chris Smaill

I deliberately did not mention the location of the discussion. Consequently, students needed to commit to the discussion in advance: they could not simply turn up on the

day and take part. Knowing how many to expect for the discussion enabled me to organise a suitably-sized venue, and adequate refreshments.

5.7.4 E-mail communications

I received my first unsolicited e-mail about OASIS even before I had drafted my research proposal. It was most complimentary, and in it the author clearly stated why he was so enthusiastic about OASIS. Since then I have received many further unsolicited e-mails concerning OASIS. Some have been enthusiastic about OASIS while others have complained about some feature of OASIS, or some practice centred on OASIS. Some of these e-mails have been very specific: for example, drawing our attention to some question for which there is an error in the answer or an ambiguity in the wording. These e-mails have been helpful in enabling us to improve OASIS and the questions delivered by OASIS. Such e-mails have been replied to promptly (almost always within a 24-hour period), acted on, and then usually deleted once the issues that they have raised have been dealt with.

A number of e-mails were also received that were more general in their content and more lasting in their relevance; for example e-mails that praised OASIS for its ability to allow students to repeat a problem with different numerical values, or complained about the way students can help each other during OASIS assignments. The value of such e-mails was quickly realised and I made an effort to encourage them, as explained in Sub-section 5.7.2. These e-mails helped flesh out a picture of how students felt about OASIS and how they used it. Often, they also helped improve OASIS or the way in which it was implemented. As such, they played an important part in the action-research programme. Such e-mails were replied to, and were then saved for future reference. The content of these e-mail messages was analysed at regular intervals using the software package N6, in the same way that the transcripts of interviews and focus-group discussions were analysed. The e-mails I received about OASIS were almost invariably focused on one issue. In this respect they were unlike interviews; however they did quite frequently provide considerable depth on that one issue. At times, constructive and elaborate explanations of ways in which OASIS could be improved were received.

5.7.5 Information collected automatically by OASIS

The software package OASIS was always seen not just as a tool to promote student learning but also as a tool to help instructors understand student learning patterns. To this end, it has always recorded considerable information about student activity, during both practice sessions and assessments. For example, when a student is working on a particular assignment, she or he may revisit one question in the assignment several times before finally submitting her or his answers to the assignment. OASIS records not only the submitted final answer to that question, but also what answer was recorded for that question each time it was revisited by the student. This practice has two benefits: first, if there is a computer crash, or if the student runs out of time before submitting his or her answers, the current answer that the student had entered for that question is not lost, and, second, student behaviour in terms of revisiting questions and changing answers can, in principle, be studied.

In reality, the situation is more complex. Although an almost overwhelming amount of information has been recorded by OASIS, it has often proved difficult to extract it from the data-base. In particular, much of the information generated by student activity prior to 2002 has been effectively ruled off-limits by a complete rewriting of the OASIS software package in a different programming language. In order to render the more recently gathered information readily accessible to instructors who did not possess the time or programming expertise to extract it from the depths of the data-base, I secured funding to hire students to write software that could be incorporated into the main OASIS programme that would make statistical information about student activity readily accessible. The first student to do this was hired to work over the summer of 2003 - 4. A second student was hired to work over the summer of 2004 - 5 and into the first semester of 2005. This second student was partly funded by the Physics Department which was also interested in accessing statistical data about student behaviour and performance. In each case, prior to the student beginning work, I initiated discussions with academic staff from my own department and also from the Department of Physics in order to determine what statistical information about student activity on OASIS would be most valuable. The features to be added to OASIS were agreed on following these discussions.

As a result of the work of these students, instructors can now readily examine study patterns at an individual or class level. For example, instructors can quickly see how student practice activity increases dramatically in the week prior to an assessment. Data for a complete set of questions can be displayed, quickly revealing which questions are being attempted most often by students: such data clearly indicate the questions that students are finding difficult. Instructors can also look at individual questions to see what percentage of students are getting each part of each question correct, and at the record of an individual student to find out how many practice questions he or she has attempted on a day-by-day basis. Much of this information is presented in graphical as well as numerical form.

The current reality is that a very large amount of information is available to instructors. For example, when a class of students completes an assignment, the individual record for each student can be readily accessed. This means that, with a few mouse-clicks, the actual assignment the student did, complete with diagrams, can be displayed, along with her or his answers, together with the correct answers. This is actually an essential feature since, in its current version, OASIS can provide assignments where different students receive different problems for each question in the assignment, and, further, each of these alternate problems is supplied with one set of numerical values chosen from many possible sets. A full record of each assignment completed by each student is therefore necessary. Consequently, to properly record an OASIS assignment completed by a first-year class that consists of 600 students entails a considerable amount of data storage.

One feature that is of interest to instructors is the length of time students spend working on each question. However, when programming work began to make these times available to instructors, it was found that the OASIS software was not recording all the relevant information: while it was recording the time and date of initial question access, it was not recording exit times for questions, and so the time spent on questions could not be calculated. This problem has now been rectified. Unfortunately, it is not possible to determine how long a student mindfully spends on a question. For example, the student may access the question, spend just a few minutes solving it, but then not submit his or her answer for quite some time. Thus, information about the time taken to solve problems derived from practice sessions

may not be very reliable. However, students should be more focused during OASIS assignments, given that there is a time limit, and data derived from this source should be more useful.

The information collected by OASIS about student activity is made available for three distinct, but not unrelated, purposes. First, it can be used by students to monitor their learning. This use applies to both the practice mode and the test mode of OASIS. In practice mode, students can monitor their own progress through the problems. They can see which questions they have attempted, how many times, and success rates on each. Test mode provides a fairly direct and timely measure of their current achievement level. The extent to which students do make use of the information provided to them can be revealed through interviews, surveys, and the like.

Second, the information provided to students, and more, is also provided to instructors who can use it to improve student learning. For example, an instructor may notice that a particular question has a low success rate, or that a particular question is being attempted an extremely large number of times by students, and decide to explain or teach to that particular question in a lecture. Alternatively, it may simply be that the question needs to be rewritten in a less ambiguous way. At another level, an instructor may notice that students are tending to leave their revision for an assignment to the last minute, or that students are typically completing their assignments well within the time allowed but are scoring poorly on them, and decide that it is time to discuss study skills or examination techniques during lectures. Thus, the information recorded by OASIS can be used in improving student performance at the course level, but can also be used to illuminate student study patterns at a very general level. Whether instructors do use the information provided by OASIS for these sorts of purposes, and the extent to which they do so, can be revealed by interviews with those instructors.

Third, the data collected by OASIS about student activity can be used to improve OASIS itself or its implementation. As an example of the former, features commonly used by students or instructors can be expanded, while features that receive little use can be investigated further in order to determine the reasons why. As an example of

the latter, server load information may indicate that holding assessments on two consecutive days for two separate large classes is to be avoided because the very large load produced by one class practising for the second assignment coincides with the load produced by the other class completing the first assignment.

5.7.6 Interviews with instructors

Interviews with instructors were on a somewhat different footing to those with students. The instructors, my colleagues, were well aware of my large interest in OASIS. This awareness certainly did not prevent them from making negative comments about OASIS (particularly those who had never been directly involved with it!) but it was considered that they might put a somewhat positive spin on their experiences with OASIS in a formal, recorded, interview situation where I was the interviewer. Fortunately, there were corroborating ways in which to gather interview data from my colleagues. In June 2003, an academic from CPD, experienced in interview techniques and knowledgeable in the field of CAL, was invited by ECE to conduct formal, recorded interviews with all staff members who were currently or had ever been significantly involved with OASIS. This academic produced a short summary, 'OASIS Impact Evaluation', based on her findings and she also made available to me her interview notes. I have myself been involved in regular, sometimes daily, informal conversations with my colleagues concerning OASIS. These conversations typically concentrated on either the good features of OASIS or ways in which it could be improved. By the end of 2005, a number of new academics had adopted OASIS, and OASIS itself and the way in which it was being implemented had undergone some significant changes. Therefore it was decided that another round of formal staff interviews was justified. These were conducted by myself, being taped, transcribed and checked in the manner previously described for student interviews.

Again, as for the student interviews, I normally began the recorded part of each interview by asking the instructor to tell me about her or his experiences with OASIS. In doing this, the aim was to give instructors the maximum freedom for their comments. However, prior to the first interview I prepared a list of questions so that I had a prompt if an instructor was not very forthcoming and I was lost for words

myself. Occasionally I did refer to the list to check I had covered the areas that I had intended to cover. The questions on the list are given below:

1. In what course do you use OASIS? What year-level is that course? How many students take that course?
2. Why did you decide to use OASIS in your course?
3. How many OASIS problems are there and what fraction of the course is covered by these problems?
4. How is OASIS used in the course? Practice? Assignments? Tests?
5. Have traditional tests or assignments been replaced by OASIS assessments? Or have extra assessments been added?
6. What percentage of the final mark is assigned to OASIS assessments? Do you feel that this percentage is about right?
7. Do you set the students both OASIS problems and textbook problems? If so, do you notice a preference for one type or the other?
8. Are there some content areas that OASIS is particularly well or particularly poorly suited to? Could you tell me more about that?
9. Are there some skill areas that OASIS is particularly well or particularly poorly suited to? Could you tell me more about that?
10. Do you notice any patterns in the ways in which students use OASIS? These patterns could involve: when they use it, which questions are attempted the most, which sorts of questions produce the most difficulties, what students do when encountering difficulties, and whether students work alone or in pairs or groups.
11. What level of student collaboration would you regard as acceptable for assignments?
12. Did you feel that the students' grasp of the subject has improved as a result of using OASIS? Could you tell me more about that?
13. What features of OASIS do you like?
14. What features of OASIS are not so good?
15. Could you please compare and contrast OASIS assignments/tests with traditional assignments/tests?
16. Would you like to go back to the pre-OASIS days? Could you tell me more about that?
17. What would be the most effective way for students to use OASIS?

These questions were used as a framework for interviews: however, they were not expected to cover and certainly did not cover the whole range of instructor responses. Most interviews departed considerably from any script suggested by the above questions. For example, one academic who previously used OASIS no longer does so as he has become a full-time, externally funded research professor and his teaching duties have been reduced virtually to zero. However, as an early adopter of OASIS, and as an ex-Head of Department, his opinion was valued. In this case the interview, as expected, did not focus on just one course but instead was more general in nature.

5.8 Action-research cycles

In Section 5.3, action research was described as evolving through a series of cycles, each entailing planning, action, observation and reflection. For this research programme, it was envisaged that each cycle would last one semester. During each semester, data would be collected via surveys, interviews, focus-group discussions, and directly by OASIS itself. Analysis of this data would then take place, primarily in the break between that semester and the next semester. After careful reflection, improvements to OASIS itself or the way in which it was used would be implemented in the next semester. This pattern certainly made sense: students could hardly be surveyed more than once about OASIS in a single course, while the breaks between semesters would provide time to analyse and reflect on collected data, to propose and debate changes in the way OASIS was being used for teaching and learning, and to write and implement changes to the software. During periods when I was lecturing, it seemed unlikely that there would be sufficient 'spare time' available for all the above research-related activities.

In reality there was a reasonable but not perfect correspondence between cycles and semesters. There were a number of reasons why the correspondence was not perfect. First, data were collected on a regular basis and, therefore, clear possibilities for improvement could present themselves at any time. For example, when the first OASIS assignment was run in 2002, I was informed that some students had worked in teams, with one student doing the assignment while the other members of the team would use the practice facility to attempt to find the numerically identical version of the current assignment question. While it was very hard for them to find the numerically identical version in practice, it was decided to disable the practice facility during assignments. This change was introduced for the next assignment, just two weeks later. Second, an urgent problem may manifest itself at any time. For example, in the second semester in 2005, a programme modification carried out to increase OASIS speed in the face of burgeoning student numbers and overloaded servers was found to introduce a serious fault and had, itself, to be modified as soon as possible. Third, some changes take longer than expected to implement. For example, programming code that is written in response to a request for a new or

changed feature in OASIS may not be produced in time for the start of the next semester and may therefore, instead, be implemented part-way through the semester.

As the above examples suggest, the notion of perfectly well-defined cycles of one-semester length did not always happen in practice. Data are collected on a continual basis, reflection can occur during the semester as well as between semesters, and it makes sense to implement beneficial changes as soon as possible. Further, changing circumstances may necessitate immediate action. The facts that courses have a duration of one semester, that courses were surveyed just once about OASIS towards the end of the semester, and that between semesters there is more time to analyse, to plan, and to rewrite software, impose some order and regularity on what would otherwise be a somewhat haphazard activity. The formal online OASIS assessments themselves impose another set of cycles, or sub-cycles. Each assessment, together with the considerable student practice activity that takes place in the period prior to the assessment, generates much data and, consequently, reflection. Changes suggested as a result of such reflection, if they can be implemented readily, are often implemented soon after the assessment in time for the next assessment. Larger changes may have to wait longer, perhaps till the next semester. Overall, the picture is one of semester cycles and assessment sub-cycles on which are superposed further activities that may have no easily discernable temporal pattern.

That the cycles are not always uniform in length or even predictable in advance should not be seen as a weakness. Rather, this is consistent with the responsive and flexible nature of action research: "In action research, standardisation defeats the purpose. The virtue of action research is its responsiveness" (Dick, 1993). Cycles need not be of the same length and, further, one cycle can start before an earlier one has finished: "There are cycles within cycles within cycles. Some extend across an entire study. Others occupy only minutes or less. The result can be a very flexible and responsive process" (Dick, 2002). Cardno (2003), speaking in a similar vein, notes the possibility of mini or 'spin-off' cycles: "Action research is a flexible and dynamic process. There is always the possibility that, as the research group begins to investigate an issue, a spin-off issue may seem important enough to be considered as a mini-cycle of action research" (p. 14).

5.9 Ethical considerations

The University of Auckland has strict guidelines that are to be followed when research involves human participants. For example, even research that involves only the use of anonymously-completed questionnaires must be approved by UoA's Human Participants Ethics Committee and the first step in gaining such approval is the completion of a five-page application form. My situation involved much more than simply an anonymous questionnaire. In addition, it involved face-to-face interviews, group discussions, and data being collected automatically for all students by the software package OASIS. The procedures I followed, outlined in this section, were guided by the requirements of the Human Participants Ethics Committee. This research programme was approved by the Committee, first in 2002 and again in 2005, each time for a three-year period.

An information sheet was prepared and given to all students. This sheet explained the nature of my research and also noted that its ultimate aim was to improve student learning. The information sheet was provided as a page in the course-book that all students received. Providing it in this way ensured that they did actually retain a copy of it and could refer to it if desired. In my first lecture with each class I explained the nature of OASIS to the students and drew their attention to the information sheet in their notes. In addition, I produced a more specific information sheet for potential interviewees, and a third information sheet for the Head of Department whose written permission I was required to have prior to embarking on this research. A permission form was also prepared for this purpose. Again, a special form was prepared for the purpose of obtaining written permission from potential interviewees prior to each interview or focus-group discussion. Copies of these permission forms and information sheets, together with a lengthy application form, were all submitted to and approved by UoA's Human Participants Ethics Committee prior to the research taking place. The text of some of these information sheets and permission forms is reproduced in Appendix two. The actual sheets and forms were, as required printed on University of Auckland letterhead.

Essentially these application and approval processes are imposed to ensure that three conditions are met. First, those involved in the research must give informed consent

and can be shown to have done so. This was achieved through the information sheets, verbal information, and the permission forms, as outlined above. All signed permission forms are required to be stored in a locked cabinet on University premises for six years. It should be noted that students were invited by email to take part in interviews. An email invitation is surely less threatening than a direct face-to-face appeal. Further, in most cases, as outlined in subsection 5.7.2, the email was addressed to the class in general, not to individual students. In this way, students could hardly feel that they were being 'put on the spot'. In fact, consent for interviews was obtained twice from all interviewees. First, prior to actually taking part in an interview, students had to reply positively to the email invitation and, second, they were asked to sign a consent form. On the consent form it was pointed out that the interviewee could ask for the tape to be turned off at any time and could ask, without giving a reason, that the information collected be withdrawn from the research at any time during the six-month period following the interview.

Second, confidentiality must be ensured at every stage of the research. In this research programme, all surveys were anonymous and students could not be identified from the forms they completed. Data are securely stored on University premises: data in paper form are stored in a locked cabinet in my own locked office, while data in electronic form are stored on the University system protected by password. Collected data will be used only for this research programme. After five years the data will be destroyed in a secure manner. For example, paper data will be shredded. When files of recorded interviews were sent to be transcribed, they were given file-names from which the participants could not be identified. Further, I made a point of never referring by name to any of the participants during the course of any interview. Each class of students using OASIS consisted of well over a hundred students, making it impossible for anyone to identify individual students, for example, on the basis of a few characteristics that came through in quoted comments.

Third, research must be conducted in such a way as to protect the participants from harm. Fortunately, the nature of this research was such that meeting the conditions of informed consent and confidentiality was essentially enough to protect the participants from harm. Using the software itself could hardly be described as a harmful experience for the students; certainly it would appear to be a more attractive

experience for them than using the textbook! It could be suggested that students who made negative comments about OASIS could disadvantage themselves when it came to test and examination marking. Only I could produce such a disadvantage, since any negative comments I passed on to other staff were passed on without attribution. In my defence I can say that I always marked as fairly and carefully as humanly possible (OASIS assessments were of course marked by the computer), that I did not compile a list in any form of people who expressed negative comments about OASIS, and that I did not re-read any such comments prior to marking. I worked hard to create a climate of trust in my classes and in the research process, and to some extent this work must have been successful, otherwise students would have been reluctant to make any criticisms at all.

5.10 Summary

In this chapter it was argued that action research does provide a way to conduct genuine research, and that it is particularly suitable for insiders in educational settings. With its dual emphases on research and action that improves the situation of those in the research setting, action research was judged to be the most appropriate research paradigm for achieving the aims of developing, implementing and validating OASIS, a Web-based software package that, through providing online practice and assessment opportunities, is designed to both improve student learning and reduce instructor workload.

Both quantitative and qualitative data were collected across a range of classes, through several one-semester research cycles, and by a number of different methods. These methods included surveys, interviews, focus-group discussions and informal conversations and email exchanges, and were augmented by data collected automatically by OASIS itself.

All the above, when combined with ethical and well-documented research practices, as well as an ongoing commitment to maintain dialogue with students, staff and other stakeholders in the research, was regarded as sufficient to ensure process validity. The case for outcome validity can be made only after the outcomes themselves are presented and discussed, in the following chapters.

Chapter 6. The OASIS Prototype: 2000 - 2002

6.1 Purpose and outline of chapter

By virtue of the action research programme, changes were regularly made to the OASIS software itself and the manner in which it was implemented. The description of the development of OASIS and its implementation begins in this chapter and continues in the following chapters. In these, the order of presentation follows the actual order of events. It was considered that this order would be more logical and would also better capture the essence of the action research process (with its planning, acting, observing and reflecting) than, for example, grouping all the student interview findings together in one chapter and devoting other chapters to other facets of the investigation. The presentation is semester by semester since each cycle of the action research programme was one semester in duration.

Because of the iterative nature of action research, it was normal for the first three research goals and their associated research questions (see Section 5.6) to be addressed in each cycle. However, initially software development was critical and therefore this chapter has a particular focus on the first goal.

The state of affairs that prevailed before my arrival at UoA is outlined in Section 6.2. My initial involvement with OASIS in the first semester of 2002 is the subject of Section 6.3. This semester also saw the first-ever implementation of OASIS assignments. During the second semester of 2002 OASIS was used for the first time with a year-one class, a class that was more than twice as large as any class that had previously used OASIS. This use of OASIS in year one, together with the lessons learned and the changes proposed, is the focus of Section 6.4. At the end of 2002 it was decided that the existing software needed to be rewritten because it did not adequately meet the first research goal's criteria. The decision to rewrite the OASIS software and the key requirements for the new version are described in Section 6.5. A summary of the chapter is presented in Section 6.6.

6.2 Prehistory: 2000 and 2001

OASIS had been created and was in use in the Department of Electrical and Computer Engineering prior to my arrival at the start of 2002. At that time, the name OASIS was an abbreviation for Online Assessment Integrated System. Developed during 2000, it was described as a Web-based tutorial and examination system and had been motivated in part by spiralling marking loads created by significantly increased class sizes, and in part by a perceived need to provide an instant feedback, formative assessment tool to help students better develop their problem-solving skills. It was believed that the skill of solving electrical circuits was to electrical engineering what the skill of performing simple 'times-table' multiplication was to mathematics (Bigdeli, Boys, Calverley, & Coghill, 2001).

Many of the broad features of the prototype version of OASIS are still present in the current version. It has always consisted of a question bank and a database that records student activity. Students access OASIS by using a Web browser. No special software is needed on the computer used by the students since all the processing is done on the server. The prototype OASIS software package was written in the PHP programming language and ran on the Linux operating system using Apache Web-server software and the MySQL relational database. Although the software package was developed over a period of some time in a somewhat ad-hoc manner in response to ongoing suggestions and requirements, it still proved reasonably fast and efficient. The original server was a mere dual 300 MHz Pentium II with 512 MB of RAM, yet it was able to handle 200 users concurrently with no loading problems.

One reason for the speed and light server loading of OASIS is found in the way it marks questions. Part of each question in the question database is a set of 200 to 300 numerically different versions of the same problem. When a question is delivered to a student, one of these versions of the question is randomly chosen. When the student subsequently submits his or her answers to the question, these answers are compared to the answers already stored in the question database for that version of the question. Marking answers therefore becomes largely a matter of comparison rather than calculation. In fact, some calculation is still involved because a submitted answer is

deemed to be correct if it is within a given tolerance of the correct answer. In the prototype version of OASIS, the tolerance was fixed at 1%.

There was a price to pay for this speed. In order to produce the set of numerically different versions for each question, some programming specific to that question had to be carried out. In addition to the time and expertise needed for this, each question required the production of a quality graphic of the relevant circuit, and the uploading of the set of numerical values, the graphic and the wording of the question itself, to the question database. It was also a non-trivial task to ensure that the right numerical values appeared in the right places by the appropriate circuit components in each graphic: while the underlying graphic was fixed for each question, the numerical values that appeared on it were not. In fact the time and expertise required to create a question and get it functioning correctly on the OASIS question database were considerable, and sufficient to stop most staff using OASIS. Consequently, senior students were hired and trained to encode questions and to upload them to the question database. Academic staff created the questions, supplied explanations of how the answers were derived, and checked that the questions, once placed on OASIS, were actually functioning correctly. Students were paid $50 per question.

The first mode offered by OASIS to students was the practice mode. In this mode, a student would log on, choose a course, then part of a course, then a particular question. Figure 6.2.1 shows a typical question from the year-three course 'Electronic Devices and Technology'. Note that there is a 'save' button that may be used to save the question and any answers submitted if the student chooses to defer completing the problem. When the 'submit' button is used, OASIS marks and records the results and supplies the student with the correct answers. The result of such a submission is shown in Figure 6.2.2. The process is illustrated in more detail with screenshots from the current version of OASIS in Appendix three.

After submitting her or his answers, the student can choose whether to attempt a numerically different version of the same question (by using the 'try again' link), or a different question. Note that on the left side of the display there is some information for the student about her or his performance to date on this and the other OASIS questions pertinent to the course.

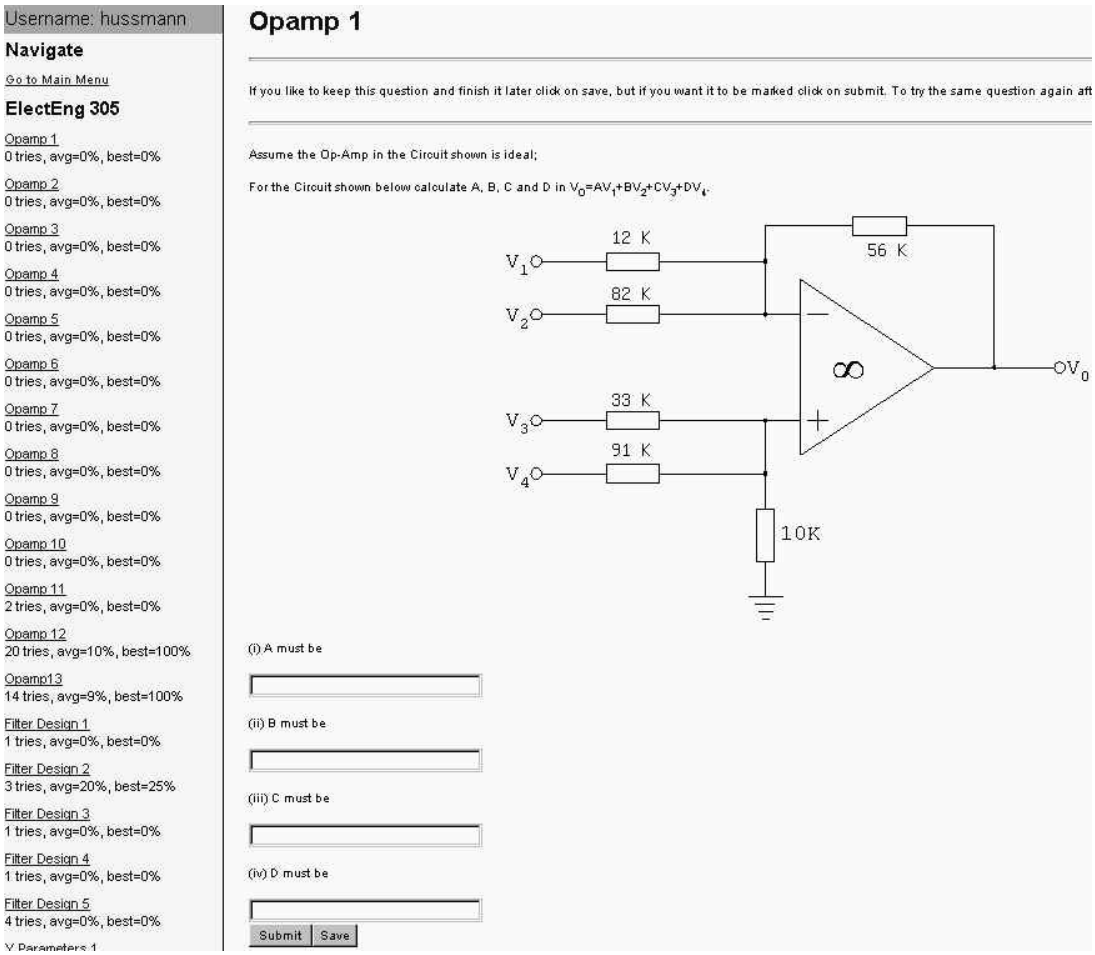
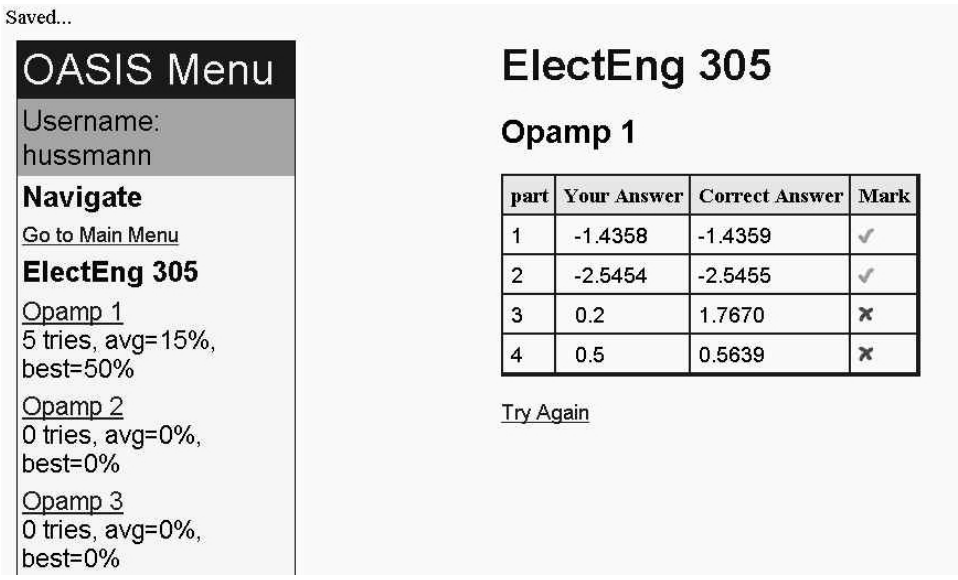


Figure 6.2.1: Year-three operational amplifier question delivered by OASIS.



part	Your Answer	Correct Answer	Mark
1	-1.4358	-1.4359	✓
2	-2.5454	-2.5455	✓
3	0.2	1.7670	✗
4	0.5	0.5639	✗

Figure 6.2.2: Illustrative result of submission of answers to an OASIS question.

Further software development enabled OASIS to operate in test mode. Screenshots of an assessment from the current version of OASIS are shown in Appendix four. In test mode, students can only log on to OASIS from supervised rooms. Tests consist of a small number of questions, perhaps six, to be done within a specified time limit. When students are satisfied that they have answered these questions to the best of their abilities, they submit their answers and receive their final marks. Concerns about the server's ability to handle large numbers of concurrent users, together with a lack of rooms that could be supervised that also contained large numbers of computers, generally led to tests being somewhat hybrid in nature, half PPA and half CBA, for example, or being done in two shifts, with the second shift directly following the first.

During 2000 and 2001, OASIS was used in three year-two courses and one year-three course. It contributed between 30% and 50% of the students' coursework in each of these courses. Coursework itself typically contributed 30% to the final course grade. Students generally found OASIS easy to use, appreciating the interactive environment and the instant feedback. There was much anecdotal evidence to support the introduction of OASIS as a learning tool. In order to produce some more concrete evidence, students in the 2001 year-two course 'Circuits & Systems' were given three weeks of exposure to OASIS practice prior to an OASIS test. Following this, a simple survey of the students was conducted. The results of this survey are summarised in Table 6.2.

Table 6.2

Student survey results for OASIS, 'Circuits & Systems', 2001

Item	Agree or strongly agree	Disagree or strongly disagree
OASIS was easy to use	48%	14%
OASIS helped improve my skill level at solving electrical circuit problems	73%	7%
I am more confident about my learning after using OASIS	51%	16%
OASIS helped me prepare for the test	68%	7%
It would be a good idea to have OASIS in other courses	51%	18%

As can be seen, the survey results are encouraging, particularly for the question concerning improved skill levels. For some reason, this returned a better result than the fourth question, even though the test consisted of questions that the students would have previously met during OASIS practice. For the same course, 'Circuits & Systems', a comparison was also made between the results that students gained in the first test in 1999 and in 2001. The tests were of similar difficulty. In 1999 students had prepared for the tests using problem sheets backed with tutorial assistance. In 2001 students prepared for the test using OASIS practice. Student achievement was significantly higher in the 2001 test, with 60% of students gaining 50% or better in the test, compared to 50% of the students in 1999. Further, in 1999 the upper quartile was under 70% while in 2001 it was over 80%.

However, while the test and survey results are encouraging, there are some good reasons why too much reliance should not be placed on them. First, the survey was fairly simple in nature, with all five items being written in a positive way about OASIS. Second, given that the test following the practice was based on the practice questions themselves, students certainly should have found OASIS to be helpful. In fact, the result that only 68% of the students considered that OASIS helped them prepare for the test is surprising. Third, the chance to practise on the actual test items before the test itself may well have lifted student performance on the test in 2001 beyond what was measured in 1999. More evidence would be needed in order to convince sceptical outsiders that OASIS really could improve student learning.

6.3 First OASIS assignments: semester one, 2002

6.3.1 Introduction: 'Circuits & Systems' and my role in it

I joined the Department of Electrical & Computer Engineering at the start of the 2002 academic year. Shortly after this, I was informed that I would be one of four teachers teaching the year-two course 'Circuits & Systems', a core course for both Electrical & Electronic Engineering and Computer-Systems Engineering. My part of the course would be the first third, and a number of Mechatronics students would also be taking this part of the course. Altogether, the enrolment would amount to around 220 students. The stated overall objective of 'Circuits & Systems' was to provide a good understanding of the way electrical circuits worked. My part of the

course covered electrical circuit components, theorems for and analysis of direct-current and alternating-current circuits, and phasor and complex-number representation of single-phase alternating-current circuits. The course was delivered over a 12-week semester with four lectures per week. No separate tutorials were held, but some lecturer and student problem-solving did regularly take place in the lectures. I also held frequent office hours. There were three compulsory laboratory exercises, each of which was assessed (pass/fail) on the basis of a student write-up and a follow-up interview. However, these did not contribute to the final grade, which was based on one three-hour examination (70%), two tests (20%), and four OASIS assignments (10%). This last involved somewhat of a departure from previous years: while an OASIS test had been used in 2001, OASIS assignments had never before been tried in any course. No other significant changes, either to the course itself or to the mode of assessment, were made to the pattern established for this course in previous years.

As outlined in the previous section, students had previously used OASIS in 'Circuits & Systems', particularly in the first third of the course, which had been taught by the then Head of Department who was an energetic supporter of OASIS. Taking over this part of the course from him, it was only natural that I would also use OASIS myself. It certainly appeared to me to be a most useful teaching and learning tool. It was only after I had worked with OASIS for some time that I decided its development and implementation would be a suitable subject for my doctoral research, more appropriate, in fact, than the subject I had previously chosen which was grounded in high-school Physics, my previous area of employment.

Once the semester started, I was kept rather busy preparing and delivering my four lectures a week, preparing course notes (which I put on the Web for students to download), reading through the two recommended textbooks to identify and check suitable problems for the students, and writing more OASIS questions to better cover my part of the course. In all of this, I was never very far ahead of the students. The advantage of this was that later parts of the course were prepared in the light of my experiences with how the students had handled earlier parts of the course.

6.3.2 Towards the first OASIS assignment

There was a clear need to produce more OASIS questions. There were two main problems with the existing set of questions. First, the coverage of topics was very uneven. For example, in the topic-area of direct-current there was a large number of questions on Thevenin equivalent circuits and relatively few questions on most other areas, while in the topic-area of alternating-current there was only one question, and it was of considerable difficulty. The second problem was the general level of question difficulty. The previous lecturer had been the Head of Department, an extremely knowledgeable and intelligent individual who had maintained a very high standard with the questions. It was my opinion that the students would struggle to master the questions already on OASIS. While being able to correctly answer these questions was an appropriate goal, I considered that some easier questions would help the students reach the demanding standards set by the existing questions. I considered that student learning would be enhanced if the students could work their way through a set of questions, starting with rather easy questions and finishing with quite demanding questions. There was another reason for starting with rather easy questions: the students would initially be unfamiliar with the OASIS software and it would make sense for them to work on questions that they felt very confident about while they were building familiarity with the software itself. In practice, there was not sufficient time for me to produce and implement more questions for the first part of my fraction of the course, the fraction that dealt with direct-current circuitry. This was due to the total length of time required for me to create the questions, for a student to be employed to encode the questions and upload them to OASIS, and for me to check that the questions appeared and functioned correctly. However, I was able to successfully add to the question set for the part of the course that dealt with alternating-current circuitry, and this was far more pressing, given that there had been only one OASIS question covering this part of the course.

Soon after the students began to work their way through the OASIS questions, I began to receive a considerable amount of informal and anecdotal feedback. Whenever I received several inquiries about any one question, I would address the issues with that question in a lecture, providing some hints about how to solve the question rather than providing the actual solution itself. Students certainly appreciated the instant feedback they received on submitting their answers to the

questions, and they also made good use of the fact that questions could be repeated with different numerical values. A number of students who did not correctly solve a question on their first attempt would repeatedly solve the question until they regarded their average score on the question as satisfactory. In this respect 80% was often regarded as a satisfactory score; thus a student who initially answered a question incorrectly would aim to subsequently answer it correctly four times in a row. I also received, from a number of sources, reports of 'OASIS races' in the computer laboratories. In such races, two or more students would start work on the same problem at the same time and the winner would be the student who first managed to solve the question correctly. The OASIS log also revealed that frequently there was a considerable group of students who were using the software up to and beyond the hour of midnight. Overall, there was a high degree of OASIS usage and an air of enthusiasm about this usage that could not be entirely contributed to the software's novelty. There certainly appeared to be far more effort put into mastering the OASIS questions than the textbook questions that I had also set. It could be argued that the students were practising the OASIS questions because they knew the assignment would be made up from those questions rather than from the textbook questions. Even so, the effort seemed quite remarkable, given that each of my two assignments would only contribute 2% towards their final grade.

On several occasions, students came to see me during my office hours and asked questions like: "I'm finding these OASIS questions quite difficult; do we really have to be able to do these questions?" My reply was always the along the lines of: "Well, that's the standard". The students would then pause for thought, nod, and accept the truth of what I had said. We would then work through one or more questions together so that they would be able to reach 'the standard'. After this had happened a few times, I realised that the OASIS questions really did provide a means of helping to define the required standard in courses. The fact that the questions were repeatable with different numerical values meant that they were much better in this respect than normal textbook questions. For example, a problem involving the calculation of the equivalent resistance of two resistances in parallel could be repeated by students until they were completely sure that they really had mastered that particular skill. By contrast, a textbook might only have a few questions requiring such a calculation,

perhaps not enough to reassure some students that they really had mastered the skill, not just memorised the answer.

6.3.3 The first OASIS assignment

The first OASIS assignment took place at the end of the third week of the semester. I had scheduled it as early in the first semester as I felt was possible. There were two reasons for this. First, I wanted students to get feedback about their performance as soon as possible. It appeared that, in a number of courses, it was quite normal for students to receive no formal feedback concerning their performance till approximately half of the course had gone by. If one purpose of the feedback was to modify student behaviour and to make it more directed towards effective learning habits, then such timing did not seem appropriate. Second, I wanted students to take the course seriously right from the start and I judged that this was more likely if there was an up-coming assessment to motivate them to do so.

There are two main differences between an OASIS test and an OASIS assignment. First, OASIS tests are only available to students on certain computers, namely those in the supervised rooms designated for the test; by contrast, students may complete an OASIS assignment on any computer they choose that has Internet access. It would also be possible to begin an assignment on one computer and complete it on another computer, provided this could be achieved within the time limit. In fact, a few students did recount tales of computer failures at home necessitating hasty drives to the nearest Internet cafe in order to complete assignments. Second, while both tests and assignments have a time limit, usually one hour, such a time limit for a test would require every student to complete the test within the same one-hour period, while, for an assignment, such a time limit would require each student to complete the assignment within one hour of their log-in time. Thus, a student could complete the assignment within any one-hour period within the designated assignment period. Initially, it was decided that this assignment period would be two days. By using a two-day period instead of a one-day period, it was considered that there would be no significant risk of server overload. Since an OASIS assignment had never been run previously, there was no information available about what temporal pattern student log-ins might follow. By spreading the log-ins over two days, it was considered that

there would be very little risk of having an excessive number of students concurrently logged on to the server.

On my arrival in the Department, I had been impressed with accounts of the OASIS software, and how it could provide skills practice, tests and assignments for students. However, while the software was user-friendly for students, it was not user-friendly for staff. For example, specialised knowledge was needed to add questions to the database and also to assemble tests and assignments from the database questions. One member of the Department's academic staff who had the requisite knowledge and expertise was designated to set up tests and assignments for the Department's courses. However, when I approached him to set up an assignment, it was discovered that the OASIS software did not in fact have the capability to run assignments. Some programming needed to be completed in order to add the appropriate functionality to OASIS. As a result, the assignment took place somewhat later than intended. Once my colleague had carried out the required programming, however, the assignment did run well with no significant technical problems.

6.3.4 Reflections and plans for future action

During the two days for which the assignment was live, OASIS practice had also been available to the students. On the face of it, this seemed to be reasonable: for example, a student who planned to do the assignment on the second day would certainly want to practise for it on the first day. However, it was noted that some students worked in pairs, and while one student was working on a particular assignment question, the other student would access the same question on OASIS practice and repeatedly submit answers to it in the hope of hitting the numerically identical version of their colleague's assignment question. While the probability of success for the above method was very low, given that each question typically had 200 to 300 numerically different versions, it was decided that this behaviour should be discouraged. Therefore the decision was made that, in the future, OASIS practice would not be available during assignments. This temporary disabling of OASIS practice was the only modification made when I ran a second assignment three weeks later with the same group of students.

Traditional, paper-based assignments, where all students receive the same numerically identical versions of the same questions, are frequently rendered virtually meaningless by large-scale copying. This situation had led the Faculty's Department of Chemical & Materials Engineering replace graded assignments with supervised tests. Such copying was not possible with OASIS assignments because different students received numerically different versions of the same question. However, it was certainly true that, in some cases, students did not work individually on assignments but worked in small groups. Clearly, in such situations there would be no point in students exchanging answers with each other; rather it appeared that the focus was on the discussion of appropriate methods for the questions. My colleagues who were involved with OASIS and I judged that such discussions provided a good way for students to advance their learning. If students chose to work in small groups and check each other's answers, then the result was that each student was doing each question more than once. Again, the consensus was that this was a good rather than a bad thing.

I received a small number of complaints from students who said that while they had been doing their assignment at the University one or more students had been watching over their shoulder and therefore finding out what particular questions were in the assignment. This was seen as giving the watching students an advantage because they now knew what particular questions were going to be in the assignment and they could go away and practise them. Of course, they still had to be able to do these questions, but they no longer had to be concerned about mastering all the other questions, in order to perform well on the assignment. This was a genuine concern in the sense that we wanted students to be able to solve all the questions rather than just the six or so questions that were in the assignment. Disabling OASIS practice for the duration of the assignment was seen as a partial solution to this concern. Reducing the assignment period from two days to a single day would also have helped in this regard; however we were still uncertain about whether the system could successfully handle the increased peak loads brought about by such a change. An analysis of the results of the first OASIS assignment did, however, reveal that students who did the assignment on the first day had, on average, scored more highly than students who did the assignment on the second day. When given this information, the students who

had voiced some concern about their assignment activities being watched by other students were somewhat mollified.

Assignments are, by their nature, less secure than tests. However, it was considered that the OASIS assignments were more secure than traditional assignments, and that the relatively small amount of mark inflation through collaboration and cheating was an acceptable price to pay for the remarkable level of student learning activity that it motivated, particularly given that a mere 2% of the final course grade was at stake: even though every lecture I gave involved student problem-solving as well as lecturer problem-solving, I still felt obliged to run two office hours each day in the period leading up to each assignment for the large numbers of students seeking help.

In general, the students did well on the two OASIS assignments, averaging 80% in the first assignment and 73% in the second. This compared most favourably with the average across the rest of the coursework (57%) and with the examination average of 55%. The examination contained two questions set by myself that were based on the content area covered by the OASIS questions I had provided for the students for practice. The first of these questions was done very well by the students and the average mark of 66% was the highest for any of the six examination questions. However, the second question produced an average mark of only 43% and this was the lowest examination question average. In hindsight, it was clear that I had set the second question at too demanding a level; my expectations, perhaps based on high student performance in the assignments, had been unrealistic.

The exercise of running the OASIS assignment highlighted further ways in which the software package was not at all user-friendly for staff. Expert knowledge was needed to achieve almost anything. For example, if a student had forgotten his or her login password, I could not reset it. Nor could I reset an assignment if a student submitted it by mistake after having done only the first question. During the course of the assignment, I could not determine who had already done the assignment, or who was currently doing it. At the conclusion of the assignment, I was not able to extract the results gained by the students. For all these things and more, I was dependent on the one colleague who had the appropriate expertise. As a result of this dependence, students had to wait quite some time before receiving official confirmation of their

assignment mark. However, as a rule, this was not of great concern to them, given that they did receive their mark immediately after submitting their answers.

It was clear that, for staff, the OASIS software needed to become much more user-friendly. For students the situation was not nearly so critical, although, even for them, there were some issues with the navigation around various parts of the software. For staff, there were three main areas where significant improvements were needed. First, staff needed to be able to access the database themselves to determine how particular students were performing, and also how well particular questions were being answered. For example, it would be useful to know which students were seldom accessing OASIS for practice, and which questions were proving particularly difficult for students. This would show up through a low success rate for a question, or through a high average time spent on the question. Second, it would be helpful if staff could actually construct assignments and tests themselves. Ideally, this would be done in conjunction with accessing the database so that knowledge of question difficulty could inform the process of choosing questions for the assessment. Third, staff would find the use of OASIS less daunting if they could themselves construct and upload questions to OASIS rather than have to brief somebody else on their requirements and also, possibly, source some funding to cover the cost of this process. In order to motivate the process of producing a more user-friendly version of OASIS, I produced some wish-lists that formed the basis for discussions involving interested colleagues and myself. As an example, a wish-list focused on staff access to the database is reproduced in Appendix six.

Concerns about peak server loading were mentioned earlier in this subsection. In fact, OASIS was running on a rather small-capacity server that was not dedicated to OASIS. The consensus was that OASIS itself did not place large demands on a server. However, if OASIS was running concurrently with other software packages then there was potential for server overload during the critical times at which OASIS was being used for assessments. Therefore, it was decided to seek funding for a server that could be dedicated to OASIS. In this way, we could have a high level of control over the server load. To this end, I successfully applied for a Teaching Improvement Grant. This grant came through in September, 2002, enabling a dedicated server to be deployed for the 2003 academic year.

6.3.5 The start of the action research programme

My initial experiences with OASIS, as outlined above, convinced me that it had the potential to become a most effective learning and assessment tool. For this reason I was keen to embark on an action research programme based on OASIS that would enable me to realise its potential and also validate its effectiveness as a learning and assessment tool. Although I had already chosen and mapped out an area for doctoral research, I considered that OASIS provided a much more exciting and relevant vehicle for research. I therefore set to work writing a candidacy proposal in which OASIS was the focus of an action research programme: this proposal was written and submitted in May, 2002.

Before I could begin the action research in a formal way, I needed approval from UoA's Human Participants Ethics Committee. Approval was needed even for research that simply involved anonymous questionnaires. My situation was more complex: I also wanted to record and analyse interviews with students, and to access and process information recorded by OASIS about the performance of each student. Further, my research would also naturally involve taking into account the marks gained by students in assessments such as traditional tests and final examinations. After some initial clarifying communication with members of the committee, I submitted a proposal at the end of May, 2002. Some minor rewriting was required and this was completed by the end of June, in time for the second semester of 2002.

For the research I had in mind, it was clear some funding would be necessary. At the very least, I would need some funding for a digital recorder and interview transcription. Software, such as N6, was also seen as being most useful for interview analysis but was not then provided by the University. Additionally, it would be helpful to have some funding for travel so that I could meet other academics and see first-hand what was happening in other universities where similar software packages were being used. To this end I applied for funding from UoA's Staff Research Fund. Applications to this fund are only accepted twice a year and, by the time I had applied to the fund and had heard that I was successful, it was the end of October and the second semester students had already completed their studies. Effectively, this meant that the funds could not be put to proper use till the start of 2003.

No formal survey dedicated to OASIS was conducted with this group of students: as explained above, the course had finished before I had gained approval to use anonymous questionnaires as a research tool. However, at the end of the course, the students did, as is standard in the Faculty of Engineering, complete a survey about the course as a whole. In completing the survey, some students did write comments about OASIS. There were two critical comments about OASIS, namely "get rid of OASIS testing" (no reason given) and "modify OASIS so that it explains the answers". In connection with this latter comment, the reader is referred to Section 3.5 for a discussion of the merits of extensive feedback. Several positive, albeit brief, favourable comments were also received: in these OASIS was described as: "good", "very educational", and "useful".

6.4 OASIS year-one implementation: semester two, 2002

6.4.1 Introduction: 'Electrical Engineering Systems' and my role in it

The first year of the four-year Bachelor of Engineering degree is a non-specialist year with students taking a range of courses that spans the main departments in the Faculty. 'Electrical Engineering Systems' is one of these courses. It is provided by the Department of Electrical & Computer Engineering and is compulsory for all first-year students in the Faculty of Engineering. In 2002, the enrolment in this course was around 550 students. Since the group was so large, it was taught in two lecture streams, 10 am and 12 pm, of about 350 and 200 students respectively. The course was delivered over a twelve-week semester with three lectures per week. Students were also timetabled to attend one tutorial each week. There were 15 separate tutorial streams. Each student was also required to complete three laboratory exercises. The final grade was based on one three-hour examination (70%), two tests (20%), and a formal laboratory report (10%). This assessment scheme was finalised at the end of 2001, before I joined the Department, and so I was unable to influence it. In particular, I was not able to introduce any OASIS assessments. However, given that this course was more than twice the size of any course in which OASIS had previously been used, the implementation of OASIS assessments would have been problematical; most likely it would have been decided to defer their implementation until a better server, or one dedicated to OASIS, had been introduced.

The course was divided into four modules, with each module being presented by a different lecturer. I was responsible for the last module, on power systems, which covered such topics as electric motors, transformers, and three-phase power generation, transmission and distribution. I also assisted the lecturer in the first module, essentially defining the scope of each lecture, checking and editing lecture resources, making study materials available for the students, and setting and marking test and examination questions. This module covered electrical circuit components, and theorems for and analysis of direct-current circuits. In some ways it was the fore-runner of the part of the year-two course 'Circuits & Systems' that I had taught in the first semester. I also held office hours during the presentation of the first and last modules, and attended one session of every lecture through the course. The reason for my involvement in the first module was that the lecturer for this module was the Deputy Vice-Chancellor (research) and his workload was already excessive without the extra commitment generated by these teaching duties. On occasion, I presented lectures on his behalf when his commitments prevented him from doing so. Overall, I presented about one third of the lectures in the course and was responsible for either delivering or overseeing about half the course. This included preparation of tutorial material and lists of relevant readings and questions from the new text.

6.4.2 OASIS practice in year one

As mentioned above, students had questions to work through for tutorials and from the text. Some of my lectures also involved the students in problem-solving activities. To supplement these questions, I wrote OASIS questions that covered the half of the course in which I had direct involvement. In the case of the first module, I was able to use some of the questions I had written for the year-two course 'Circuits & Systems'. The end-result was a set of about 40 questions for the first module and a further 25 questions for the last module. Because each OASIS question contained 200 or 300 numerically different versions, far fewer questions were needed than would normally be the case. For example, in high-school mathematics, a single OASIS question would be able to replace a whole page of traditional questions on solving quadratic equations. In fact, I had been limited largely by my imagination: I would have written more questions for the last module if I had been able to think of any worthwhile questions, apart from purely factual recall questions, that varied significantly from the 25 that I had already written.

There was certainly some concern about whether students would spend appreciable time solving the OASIS questions. First, the students already had quite a collection of questions to be done for the weekly tutorials as well as further textbook questions (with answers provided). Second, OASIS was to them a completely new piece of software; consequently there would be some learning involved before they would be able to comfortably use it. Third, while a sizable proportion of the class would be continuing in the Department of Electrical & Computer Engineering, the reality was that the majority of the students would be continuing their education in other departments. These latter students might not be keen to practise yet another set of electrical engineering problems, or to master yet another piece of new software.

Students were introduced to OASIS in the first week of the course and, following this introduction, it soon became apparent that the majority of students were spending a significant amount of time using it. A lot of the evidence was anecdotal: for example, a walk through the computer laboratories would reveal a large number of students using the software, while a check of who was logged on to OASIS at various times indicated that it appeared reasonably normal for 30 to 40 students to be logged on well after midnight if there was a relevant (albeit traditional paper-and-pencil) test the next day. Based on the student inquiries I received during the semester, it seemed that the weekly tutorial questions were receiving the most attention, followed by the OASIS questions. Textbook questions appeared to be receiving the least attention.

6.4.3 Some statistical explorations

As previously mentioned, the prototype OASIS software would not yield information to academic staff, such as myself, who did not possess an intimate understanding of its data structures, and I was not able to obtain significant data about student practice on OASIS until late February the following year. When I finally obtained this data, it did support the observation that students had made significant use of OASIS, even though there had been no particular motivation (in the form of OASIS assessments, for example) for them to do so. Of the 530 students who completed the course, 415 had used OASIS for practice - 78% of the cohort. On average, the students who did use OASIS had submitted somewhat in excess of 100 answers to OASIS questions. For the full group of 530, the average student had made 80 submissions of answers to

OASIS questions. These figures were most pleasing, particularly given the fact that a number of OASIS questions contained three or four parts, and that a submission to a four-part question, for example, was counted as a single submission.

I was keen to investigate the usage of OASIS in order to illuminate two related issues. First, I wanted to find if there was a link between a student's OASIS activity and her or his subsequent achievement. Second, I wanted to find out whether it would be possible to use OASIS activity as a means of predicting future student success. One variation on this latter theme was the possibility of identifying at-risk students from the way in which they were using OASIS. To this end I looked at the 29 students who had completed the course with a grade of D+. Such a grade corresponds to a mark between 45% and 49%. These students had used OASIS somewhat less than the average student. 28% had not used OASIS at all, compared to the overall class figure of 22%. On average, students in this group had made 67 submissions to OASIS questions, compared to the figure of 80 for the entire cohort. Perhaps of more promise was the figure that a full 45% of the students had made fewer than 20 submissions to OASIS questions. It seemed likely that a student who had made such a small number of submissions had probably performed only exploratory investigations of OASIS and had not used it in any determined fashion. Perhaps, in the future, it might prove beneficial to contact students who were exhibiting such a low usage figure and warn them that they were at risk of inadequate achievement in the course. With this in mind I extended the investigation to include the 20 students who had finished the course with a D grade. Such a grade corresponds to a mark between 40% and 44%. Surprisingly, these 20 students had used OASIS more than the group of 29 students who had gained a D+ grade. Twenty five percent had not used OASIS at all, compared to the D+ cohort figure of 28%. On average, students in this group had made 76 submissions to OASIS questions, virtually matching the figure of 80 for the entire cohort, and surpassing the figure of 67 for the D+ group. The figure of 35% for the percentage of students who had made fewer than 20 submissions to OASIS questions was an improvement on the 45% figure returned by the D+ group. On the evidence provided by the above simple analysis of the submission figures of the combined group of 49 students, it seemed likely that any productive search for a link between OASIS usage and achievement

would probably need to go well beyond the statistic of the number of submissions made to OASIS questions.

Figure 6.4 presents a scatter-graph showing the final examination mark and the number of submissions to OASIS questions for those students who completed the year-one 'Electrical Engineering Systems' course in 2002. This graph does not display data for those students who made no submissions to OASIS questions. It also excludes one student who made 1025 submissions to OASIS (and who scored 71% in the final examination). The pattern of Figure 6.4 suggests that no strong relationship existed between the final examination result and the number of submissions to OASIS questions.

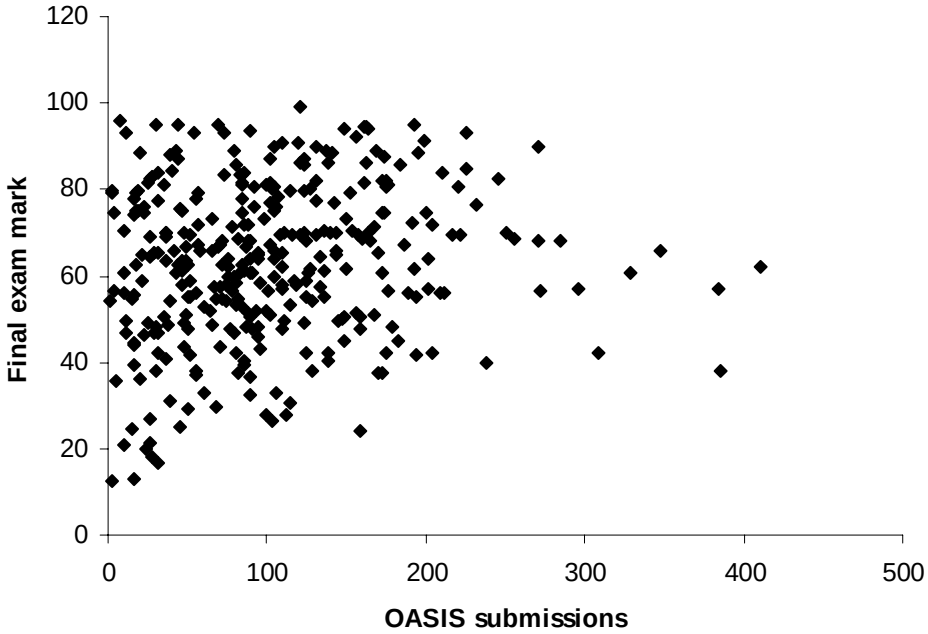


Figure 6.4: Scatter-graph for year-one 'Electrical Engineering Systems', 2002, showing relationship between final examination mark and number of submissions to OASIS questions.

In order to investigate in more detail the extent of any such relationship, the software package 'Statistical Package for the Social Sciences' (SPSS, 2006) was used to perform a statistical analysis of the data. The output from the SPSS analyses discussed in this subsection is provided in Appendix seven. For practical information on the use of the software itself, I found Weinberg and Abramowitz (2002) helpful, while I consulted Lewis-Beck (1993) for background information of a more

theoretical nature concerning regression analysis. The group analysed was a subset of the full year-one class, namely just those students for whom UEBS examination results could be obtained. The UEBS examinations were the national external examinations that students sat at the end of their final year of high school (they have since been replaced by the NCEA Level-three examinations). There were 327 students for whom UEBS examination data was obtainable. The first analysis involved a simple linear regression model with the predictor, or independent variable, being the number of submissions to OASIS questions, and the dependent variable being the final examination result in Electrical Engineering Systems. The SPSS analysis of the data produced the following equation for predicting the final examination result (EES) from the number of submissions to OASIS questions (OASIS):

$$\text{Predicted EES} = 0.058 \times \text{OASIS} + 57.7$$

The above equation predicts, for example, that a student who made 100 submissions to OASIS questions would gain a mark of $0.058 \times 100 + 57.7 = 63.5\%$ in the final examination, whereas a student who made no submissions to OASIS questions would gain a mark of 57.7%. The coefficient 0.058 in the above formula was found to be statistically significantly different from zero ($t > 3.5$, $p < 0.0005$). However, the analysis also showed that only about 4% of the final examination result could be explained in terms of the number of submissions to OASIS questions ($R^2 = 0.037$). This low percentage is consistent with the pattern displayed by the scatter-graph, which suggests that the total number of submissions to OASIS questions is not an important factor in predicting the final examination mark.

The software package SPSS was then used to perform another statistical analysis of the data using the same group of 327 students. This second analysis involved a multiple regression model with the predictors being the number of submissions to OASIS questions and the UEBS examinations total. This total is of the marks gained by the students in the five subjects they had presented for the UEBS examinations. For this total, the maximum possible score was 480. As before, the dependent variable in the analysis was the final examination result in Electrical Engineering Systems. The analysis produced the following equation for predicting the final examination result (EES) from the UEBS examinations total (UEBS) and the number of submissions to OASIS questions (OASIS)

$$\text{Predicted EES} = 0.239 \times \text{UEBS} + 0.042 \times \text{OASIS} - 26.7$$

The above equation predicts, for example, that a student who scored the same UEBS total as another student but who made an extra 100 submissions to OASIS questions would gain a mark 4.2% higher in the final (EES) examination. This small improvement in final mark would most likely be seen by both students and staff as a disappointingly low return on the extra work involved. As a second example, the equation above predicts that a student who made the same number of submissions to OASIS questions as another student but who gained, on average, 5% higher in each of his or her 5 UEBS examinations would gain a mark 6% higher in the final (EES) examination. Both coefficients in the above formula were statistically significantly different from zero (for 0.239, $t > 13$ and $p < 0.0005$; while for 0.042, $t > 3.2$ and $p < 0.001$). However, the analysis also showed that only about 37% of the final examination result could be explained in terms of the UEBS examinations mark and the number of submissions to OASIS questions ($R^2 = 0.372$). The standardised coefficients (0.581 for the UEBS examinations mark, and 0.142 for the number of submissions to OASIS questions) showed that the UEBS examinations mark was a significantly stronger contributor to the predicted final course mark. The evidence provided by this analysis suggests that the number of submissions made to OASIS questions was not a particularly useful predictor of achievement in the course. Some reflective comments are made on this matter in Subsection 6.4.5.

6.4.4 Written student feedback

No formal survey dedicated to OASIS was conducted with this group of students since the software had been provided as just one resource for practice, together with tutorial and textbook questions. However, at the end of the course the students were asked to complete a survey about the course as a whole. In completing the survey, a number of students did write comments about OASIS. Most comments were brief, perhaps partly because the survey form contained relatively little space for free responses. There were no uncomplimentary comments. Several students described OASIS as “cool”, “good” or “very good”. It was said to have “helped heaps”. For one student, it “helped gain better understanding”. Another commented similarly: “OASIS is an excellent resource for practice and understanding”. Perhaps feeling confident of success, one student wrote: “Thank goodness for OASIS, Hambley [the prescribed text] and some work”.

Although I had solicited no feedback about OASIS in this course, I did receive a few emails from students containing positive feedback. The one reproduced below in its entirety (apart from one omitted name) remains among my favourites.

> Hello Sir,
>
> I'd just like to say that the only reason I passed the ELECTENG101 test (23/30) was OASIS. I couldn't fully comprehend what was going on in the lectures; the tutorials were, to be frank, utterly useless; and even going to [a staff member] didn't help as much as OASIS. The ability to do questions repetitively until I was sure I knew exactly how to do them was 'the best thing since sliced bread'. Of course, I did have to have a certain level of conceptual knowledge to do OASIS, so for those that were totally 'blind' in regards to the concepts, it wouldn't have been all that much help.
>
> There's just one request I'd like to make: please put up similar questions for the rest of Module 1 and further modules that we do. I'm having real difficulty understanding what is going on at the moment with regards to Module 2.
>
> Thank you very much for OASIS.
>
> [email from a year-one student, 13/09/2002]

The above request for more topics to be covered by OASIS was to become the most common request in feedback concerning improvements to OASIS.

6.4.5 Reflections and plans for future action

It was clear that, even without the motivation of OASIS assignments, most students had used OASIS extensively. Although only approximately half the course was covered by OASIS questions, the average student still made about 80 submissions of answers to OASIS questions. There was a wealth of anecdotal evidence to suggest that students found it easy to use and that they believed it helped them improve their understanding. The examination results also supported this notion. The first and last modules of the course were supplemented by OASIS questions for the first time in 2002. In 2002, 16% of the students failed the examination question that tested students on the material covered in the final module. This percentage was a significant improvement on the previous year's figure (34%), even though the 2002 examination question was, if anything, more demanding than the 2001 question. Similarly, in 2002, 31% of the students failed the examination question that tested students on the material covered in the first module. The corresponding failure rate

for 2001 had been a staggering 80%. Here the direct comparison between the 2001 and 2002 situations was harder to make because there had been some issues in 2001 concerning the lecturer who had delivered the first module. In fact, this was the reason that the Deputy Vice-Chancellor (Research) had been approached to deliver that module in 2002. That aside, it was certainly clear that, in the examination question based on the first module, the students had exhibited a much higher level of skills and understanding in 2002 than in 2001. Overall, the evidence for increased student learning made a strong case for continuing and expanding the implementation of OASIS.

However, given the weakness of the correlation between the number of submissions and the final examination result (as presented in Subsection 6.4.3), it was clear that this student learning could not be measured simply by counting the number of OASIS submissions made by the students. Of course, there are a number of reasons for a lack of significant correlation. For example, students may click the submit button without filling in any answers if they wish to find out the correct answers but don't wish to attempt a solution themselves at that time. Students who simply wish to explore the functionality of OASIS and learn how it works may also do this. By contrast, another student may work mindfully for 30 minutes or more on one question, perhaps even checking the answer by re-solving the question using an alternative method, before submitting her or his answer. A simple totalling of submissions could not differentiate between these very different sorts of behaviours. Any genuinely informative statistical analysis would clearly need to be quite sophisticated. For example, rather than attempting to rate students on the number of OASIS submissions made, it might be more illuminating to look at the following: the number of right answers submitted, the number of different questions that have been answered correctly, the number of different questions that have been answered correctly on the first or second attempt, the amount of time spent logged on to OASIS, or the number of times logged on to OASIS for 15 minutes or more.

As a result of my experiences with OASIS in 2002, I was determined to continue and expand on its use in 2003. I considered that it both promoted and assessed student learning effectively. OASIS motivated students, provided prompt feedback, yet did not add greatly to the workload of instructors. I believed that it also had the potential

to provide much useful information for instructors about student learning activity, and to be a most useful vehicle for research into student learning, provided the data collected and stored by OASIS could be readily accessed and presented. However, to make such research practicable, given the difficulties faced by instructors in extracting information from the current OASIS database, a major reworking of the software was required. Therefore, I saw three main priorities at the end of 2002: adding to the question data-base, obtaining an adequate server that would enable OASIS assignments to be held with the large year-one class over a single 12-hour period, and rewriting the software to enable instructors and researchers to readily access the wealth of information stored on student learning activity by OASIS. This last priority is the focus of the following section.

Reflections on the research itself, as distinct from the software development and implementation, revealed both strengths and weaknesses. In terms of outcome validity, I considered that the learning of the students had been enhanced, and that there was a good body of supporting evidence, such as examination results and confirming comments from both staff and students. This evidence would, I believed, be augmented by further cycles of action research and data collection. The issue of staff workload reduction was more problematical. While it was clear that OASIS assignments demanded very little staff time compared to traditional assignments, I had spent considerable time writing OASIS questions through the year. It was not clear whether this question-writing activity would be merely a one-off activity associated with the setup of a new practice.

In terms of process validity, I believed that I had made a sound start to the research but that more cycles and more triangulation would be needed to establish sufficient rigour. Recently-obtained University approval would enable the surveying and interviewing of students in 2003 onwards. Up till now, only informal interviews with staff and students had taken place, and survey comments on OASIS had been gleaned from course surveys rather than surveys dedicated to OASIS. While the great majority of early results had been encouraging, interviews and dedicated surveys would greatly increase the rigour and therefore the process validity of the research. Another priority was to increase triangulation by having someone other than myself conduct and analyse some interviews.

Some degree only of democratic validity had been achieved. Both staff and students had certainly been kept up to date with developments (students through lecture demonstrations and announcements, staff through seminars and informal talks), but students had not been provided with clear opportunities to influence the development and implementation of OASIS. Although informal conversations and general course surveys had provided the more outgoing and eloquent students with some opportunities to exert influence, it was hoped that future interviews and dedicated surveys would provide a larger group of students with better opportunities.

6.5 The rewriting of the OASIS software

As mentioned previously, the OASIS prototype had several shortcomings and did not satisfactorily meet the criteria associated with the first research goal (see Subsection 5.6.1). While students found the software easy to use and supportive of their learning activities, staff were less well served. Staff could not, unless they had expert knowledge, put questions on OASIS, set up assessments, access the marks that students scored in assessments, or access data about student performance. In fact, staff needed help to execute almost any task related to OASIS. Not only was this a barrier to staff adoption of OASIS, it also placed a large workload on the one staff-member who did have the necessary expert knowledge of the software.

Instructors were also proposing some new question types that they wanted OASIS to support. For example, interest had been shown in using OASIS to deliver questions that required students to prioritise items in a list. As a simple illustration, students could be asked to rank a list of cities from most northerly to most southerly. While there is just one completely correct answer to such a question, it would be reasonable to award partial credit to answers that showed some merit. Certainly it would be unreasonable to assign zero marks to every answer that was not completely correct. Some answers would be better than others and should be rewarded accordingly. This sort of question is difficult for a human marker to mark correctly and rapidly; for this reason, computer marking seems most suitable. The term ‘fuzzy marking’ was coined in conjunction with this sort of question. Unfortunately, attempts to extend the prototype to include new question types such as this were unsuccessful.

There were other problems appearing with the prototype software. It was not particularly well designed for the heavier usage it was now receiving. As more questions, and sets of questions, were added to the question database, navigation around the system became more awkward. Further, with more classes using the software, another problem emerged: it was difficult to extract information about student performance on a per-class basis. Investigation revealed that these problems could not be fixed by superficial changes to the software.

Given the many shortcomings now apparent, it was decided to completely rewrite the software, while retaining or improving all its many good features. Following this decision, a list of requirements for the new version of OASIS was drawn up. These requirements essentially provided detail and also made some additions to the original criteria associated with the first research goal. Some related to established features or practices that we wished to retain. The requirements are itemised below.

1. Students should be able to access the system using a reasonably modern version of any of the common Web browsers (such as Internet Explorer, Netscape Navigator, and Mozilla Firefox). All common hardware and operating system configurations (such as Windows, Linux, and Mac OS) should be supported.
2. Any software that the OASIS software depends on for its operation should be free of cost and significant distribution limitations.
3. The new system should be able to use the large bank of questions that had been created previously for use in the prototype.
4. Whereas the prototype supported just one type of question with a fixed marking tolerance of 1%, the new system should support this question type and additional question types (such as those involving 'fuzzy marking') and also allow the tolerance to be varied on a question-by-question basis. For example, a question based on reading data from a graph should allow for a higher tolerance in the answer than a question based on a calculation.
5. The new system should provide both tests and assignments with a configurable time limit (in the prototype version, the time was fixed at 60 minutes). There should also be an unobtrusive remaining-time indicator on the screen. Instructors should be able to place at the start of each assessment a set of instructions specific to that assessment. Students should also be able to browse the questions

and revisit questions and alter their answers to the questions in the same sort of way that can normally be done with traditional pen and paper assessments. Since more than one assessment could be offered by OASIS at the one time, students would need to receive clear instructions backed by unambiguous navigational routes through the software.

6. Academic staff must be able to perform most of the administration tasks needed to provide practice and assessment opportunities in their own courses. In particular, staff should be able to set up assessments for their own students and extract the marks afterwards in a convenient format.
7. In connection with the above, OASIS needs to recognise which students are enrolled in what courses. Otherwise, it becomes very difficult to identify those students who should have done an assessment but did not. Since enrolment information can change on a daily basis, some regular, perhaps automated, updating of this enrolment information in OASIS would be desirable.
8. More generally, as many tasks as possible should be automated.
9. OASIS provides an excellent vehicle for research into student learning and CBA. To support such research, OASIS should record data about many aspects of student on-line behaviour and make the data readily available to researchers.
10. The prototype required a high level of skill from those individuals charged with the task of encoding questions. This task must be made as easy as possible. Ideally, it should be possible for instructors with no special expertise to enter questions into the OASIS database. The new system should still allow the use of the several utilities and methods developed for and used with the prototype for the creation of question templates.
11. The new system should be able to support a large number of registered users, and a large number of concurrent users. It would be sensible to aim to support around 500 concurrent users: although some classes exceed this number, the number of concurrent users is limited by the number of computers that could realistically be booked at one time for a supervised test.
12. OASIS should be readily scalable to allow for future development and expansion. A high degree of flexibility and scalability would be needed: first, such expansion could involve question types and marking schemes not under current consideration; second, further development could involve the work of individuals not currently involved with OASIS; and, third, such expansion could

involve other departments both within and outside the Faculty of Engineering, and could involve network and server issues as well as software issues.

13. Documentation should be produced to cover OASIS from three perspectives. First, there should be a users' manual which would be suitable for instructors using the software. Part of this manual could also serve as a user's guide for students. Second, there should be thorough documentation of the software. This documentation would prove invaluable if the author or authors of the software package left the University or were overseas when the software needed further modification, or if a critical software problem suddenly made itself apparent. Third, if the procedure of question-creating was still to be a specialised one, as seemed likely, documentation detailing this task would be most useful.

In attempting to meet the above requirements, it would need to be remembered that the primary purpose of the software was to improve student learning, and that a secondary purpose was to save instructor time.

The above requirements, somewhat theoretical in nature, were complemented by some more practical requirements in the form of financial and time constraints. The previous Head of Department, an early adopter and supporter of OASIS, had become a full-time research professor in mid-2002. The incoming Head of Department needed to be convinced of the worth of OASIS before supporting its operation and development through departmental funding and staff hours. Fortunately, an account of the implementation of OASIS assignments in the year-two course and OASIS practice in the year-one course did lead him to decide that OASIS should continue to be supported. Given the need for a complete re-writing of the OASIS software, as well as a new server, this was not a decision to be taken lightly. A colleague with particular skills in programming was assigned to the task of re-writing the software, and it was envisaged that this would take about 500 hours. Although I had some input into the specification of the new software package (as outlined in Subsection 6.3.4, for example), I had no direct involvement in its actual writing. Therefore what follows is a brief account only of the creation of the new software package, covering just the main design features and decisions.

Debian Linux had proven to be a robust and flexible operating system for the prototype and so it was retained for the new version. The prototype OASIS had been

implemented in the PHP language, and it was judged that this language itself was, in part, responsible for the difficulties in extending the prototype. For this reason, a new programming language was chosen that had been designed with larger applications in mind: Python. The prototype had successfully used the MySQL database; however, this database had sometimes proved a little unreliable, suffering occasional data corruption. Furthermore, its lack of some modern database features had necessitated some complex software code. A more powerful database, PostgreSQL, was therefore chosen for the new version of OASIS. This new database brought with it many features designed to improve performance and data integrity. Some of these new features enabled more processing to take place in the database itself, taking some pressure off the OASIS code. The JPEG and PNG formats were chosen for embedded graphics, rather than GIF; although GIF is widely-used, it contains patented components that require licensing fees to be paid.

Perhaps the biggest technical problem facing OASIS was the variation in computer environments at the client end. Among the end-users there was a wide range of Web-browsers and operating systems. To gather some information about this, one million accesses to the department's Website were logged from November 2002 to February 2003. It was found that around two-thirds of the accesses involved Internet Explorer, with approximately one-sixth using Netscape Navigator and one-sixth using neither of these Web browsers. Clearly, OASIS would need to do more than just support the two most popular browsers. In reality, even supporting just the two most popular browsers would be difficult. For example, Netscape Navigator 4 and Netscape Navigator 6 are actually completely different browsers, as are Internet Explorer for Windows and Internet Explorer for Macintosh. Further, both Netscape Navigator 4 and Internet Explorer 4 contained many bugs that would make it difficult for OASIS to run well on them. Although there was no question that OASIS would need to be as compatible as possible with as many browsers as possible, the reality was that achieving this was going to be very difficult. In fact, while the new version of OASIS was functional under Netscape Navigator 4, the screen display was most untidy.

Similarly, it was found that around one-third of the accesses involved the Windows 2000 operating system, with Windows NT and Windows XP each responsible for about one-sixth of the accesses. However, around 6% of accesses were from

machines running non-Windows systems. Therefore, it would be unacceptable to limit clients to Windows-based systems. This would, using the data above, disenfranchise over 30 students in the year-one class alone. Fortunately, the new version of OASIS was able to effectively support all operating systems.

The new version of OASIS was written with expansion and the addition of new features in mind: using simple protocols it is able to interface with external software packages so as to allow the testing and addition of new features without affecting the core functionality of the software.

The requirement to keep updated records of students and their courses was met by synchronising the OASIS student database each night with the Department's database. However, it was found that occasionally the Department's database did not contain the record for every student enrolled in a given course. This mis-match would be detected when assessment data was uploaded to the university's database, after which a manual search would need to be conducted for the assessment records of any students missing from the OASIS student database. The student's assessment data would still be recorded, but the student would not appear on the class list because the Department database did not place the student in that course.

The new version of OASIS was first tested with a class in March 2003: more will be said about its operation in the following chapter, Chapter 7, which covers that period.

6.6 Summary

This chapter outlined the three years of use of the prototype version of OASIS; the two years before I joined the Department and the first year of my employment in the Department. In this time, the considerable amount of anecdotal evidence gathered suggested that students did use OASIS extensively and that their learning improved as a result. There was also some quantitative data to support this conclusion.

However, in 2002, with more courses and students using OASIS, and with increasing requests for new features, it became clear that OASIS could not satisfy the demands

placed on it. While students found it easy to use, academic staff needed help to add questions, set up assessments, and extract data, including assessment results.

Therefore, towards the end of 2002, it was decided to completely rewrite the OASIS software. A package was needed that was more user-friendly for instructors, could handle larger numbers of students and classes, could support the many new features and question-types being proposed, and was sufficiently flexible to be readily extended and modified without the need for dangerously complex programming. It was also important that OASIS support a wide range of client hardware and software.

Based on the above, a set of requirements was drawn up, and the Department assigned an academic to rewrite the software accordingly. The new software went live in April 2003. The following chapter contains an account of the initial testing of the beta version and the first year of implementation of the new version of OASIS.

Chapter 7. OASIS comes of age: 2003

7.1 Purpose and outline of chapter

The previous chapter outlined the implementation of the OASIS prototype, 2000 to 2002. I joined UoA at the start of 2002 and so was involved for only that year. This chapter describes the events of 2003. As this was the first year of the implementation of the new version of OASIS, there was initially quite a focus on the new software and the extent to which it met the expanded criteria for the first research goal. These criteria were presented in Section 6.5, and the extent to which the new version satisfied them is discussed in Section 7.2. Hardware issues provide the subject of Section 7.3, in particular the new server that would be dedicated to OASIS for the next three years (ECE upgrades computer hardware on a three-year cycle). The simulation testing of the new software on this server is also described. The first trial of the new version of OASIS took place in semester one, 2003. This trial, effectively of a beta-version of OASIS, is the subject of Section 7.4. The success of the trial, and the increasing interest being shown in OASIS, motivated the Head of Department to appoint a committee, with myself as Convenor, to oversee the operation and management of all activities associated with OASIS. The setting-up and the terms of reference of this Committee are outlined in Section 7.5. The events of the second semester of 2003, most notably the implementation of the new version of OASIS with the very large first-year class, are described and reflected upon in Section 7.6. With 550 students, this class provided a rigorous test for OASIS assessments.

The second research goal, focused on instructor usage of OASIS, was addressed through staff interviews, the subject of Section 7.8. The third research goal, focused on the enhancement of student learning, was fully addressed in 2003. A survey dedicated to OASIS was developed and completed by both second-year students (see Subsection 7.4.3) and first-year students (see Subsection 7.6.3). Interviews with both these student groups were also conducted. These interviews are described and analysed in Section 7.7. Reflections on the year's events and plans for the next cycles of action research are presented in Section 7.9. Finally, the chapter is summarised in Section 7.10.

7.2 The new version of OASIS

The prototype version of OASIS needed to be rewritten because it could not cope with the demands of increasing numbers of classes, students and questions. The hope was that the new version would handle these demands and also be sufficiently flexible and scalable to handle future demands. Additionally, it was noted that, while rewriting the software, it would be sensible to add in certain new, desirable features. However, even just rewriting the software in a new language and attempting to make it flexible, scalable, and future-proof, was a huge task in itself. For this reason, a number of new features were not written into the software at this time. However, the new software was amenable to the future addition of these features. In what follows the new version of OASIS is referred to simply as OASIS.

In terms of the requirements listed in Section 6.5, OASIS did support all common hardware and operating system configurations. Students could also access OASIS with any reasonably modern version of a common Web-browser, although Netscape Navigator 4 did provide an untidy screen display. For its operation, OASIS did depend only on software that was free of cost and significant distribution limitations.

There was no problem in using the large bank of questions that had been created previously and used with the prototype. Unfortunately, the new database was far removed in language and structure from the prototype's database, and it was found that OASIS could not access the student-performance data that had been collected by the OASIS prototype and stored in its database. As a result of this, most of the data from the three-year period 2000 to 2002 was effectively lost.

OASIS now supported a number of new question types; in fact more types than most instructors were prepared to use. The time limit for assessments was now variable. During assessments, students were presented with a remaining-time indicator. Students sitting assessments received clear instructions and could revisit questions in the way that they could with traditional pen and paper assessments.

Navigation around the software was easy. In particular, the assessment function of OASIS was made quite distinct from the practice function of OASIS. It would be

very difficult for students to confuse one with the other, and impossible to follow a link out of an assessment and consequently lose the assessment.

OASIS was well set up to support student learning, and therefore the main objective of rewriting the software had been achieved. Unfortunately, the task took far more time than expected, and had also been held back for various reasons: for example, the late arrival of the new server. As a consequence, at the start of 2003, there was little direct improvement in the situation for instructors. Academic staff could still not set their own assignments or extract assessment or practice data from the database. Adding questions to the question database was still sufficiently complex to discourage most staff members. Staff still needed help with these tasks, although some progress had been made: for example, OASIS did now know which students were enrolled in what courses and this information was updated on a daily basis from the Department's database. As a result, even though it still required specialised knowledge, it was now much easier to extract practice and assessment data for a class, and it was much easier to add in the functionality that would permit staff members to do this themselves at some future date.

Documentation was another area needing attention. At this time, the programmer was extremely busy actually producing the new code and had little time for documentation. Further, the software was in a state of flux, and any documentation would have required constant updates to avoid becoming obsolete. However, given that the programming had been carried out by one individual, and that this programming was not documented, there was a clear vulnerability issue. When OASIS had played only a minor role in ECE, this vulnerability was not so critical; now, with OASIS more widely adopted, the issue was becoming more pertinent.

7.3 The new OASIS server

The original server used with OASIS was based on a dual 300 MHz CPU with 512 MB of RAM and an 8 GB hard drive. This server was of low specification and was not dedicated to OASIS, so problems could arise under heavy OASIS use, or when heavy OASIS use coincided with heavy use from another source. Therefore a new server was sought and obtained, funded from a Teaching Improvement Grant that I

received at the end of 2002. This new server, to be dedicated to OASIS, utilised a dual Xeon 1.8 GHz CPU with 512 MB of RAM and two 36 GB hard drives, these being configured in RAID 1 so that each replicated the other: if one drive failed, the other could take over the system without noticeable interruption or loss of data.

Some steps were taken to test the software prior to using it with classes. It was thought that the heaviest load occurs when many students simultaneously log on to OASIS and navigate to the first question, such as would happen at the start of a test. Using the benchmarking software Siege (2006) to simulate this, it was found that the server could handle up to 500 users very comfortably, but that the performance rapidly degraded as more users were added, becoming very slow for 550 users. The user database was populated with 50 000 student records and no drop in performance was measured. With no persistent faults being revealed in the software, and good evidence for the system's ability to function properly under heavy load, the decision was made to run a trial with a class. The class chosen was Circuits & Systems, and the trial is described in the next section.

7.4 The OASIS trial: Semester one, 2003

7.4.1 Introduction: Changes from 2002

At the start of the 2003 academic year, it was decided to run a trial of the beta-version of OASIS with one class, namely the year-two 'Circuits & Systems' class. All other classes used the prototype version of OASIS through semester one. As previously described in Section 6.3, OASIS had been used with the 'Circuits & Systems' class in semester one, 2002. For 2003, there were no major changes in the course itself, nor in the way it was taught and assessed. As in 2002, a number of Mechatronics students also took my part of the course. However, the 2003 trial had three small advantages over the previous year's implementation of the prototype with the class. First, the students had already gained some familiarity with OASIS, having used it for practice the previous year in the year-one course 'Electrical Engineering Systems'. Second, since the software now existed to support OASIS assignments, it was possible to run them at the appropriate time in the semester. The previous year they had been delayed one week while programming to enable OASIS to support assignments had been hastily carried out. Third, since the OASIS server was now a

more powerful server dedicated to OASIS, it was decided to run the assignments over a single 12-hour period, rather than over two days as previously.

7.4.2 The OASIS assignments

As in 2002, I lectured the first third of the course. There were four OASIS assignments during the course and I was responsible for two of them, each worth 2% of the final grade. The first assignment took place on Wednesday of the second week of the semester while the second assignment took place on Tuesday of the fourth week of the semester. Compared to other courses, these assessments took place relatively early in the course. In fact, it was highly likely that students in this course received two instances of feedback (from these assignments) before they received any feedback in any other course. There were three reasons why I had chosen to implement such early assignments. First, I did not want assessment of my work to take place when lectures in the course were being given by another lecturer: I felt that such assessments would detract from those lectures. Second, I wanted students to receive early, meaningful feedback about their learning. Third, I wanted students to make a good effort to grasp the material right from the first lecture, and I considered that an assessment (even one worth only 2%) would help in motivating the students to make such an effort.

There were no major problems in the running of these two assignments. The new, more powerful, dedicated server admirably handled the increased load produced by the shortened assignment time. There were, however, a few minor problems related to the software. For example, OASIS had been programmed to provide a warning pop-up panel which would appear when a student had just three minutes remaining in which to complete his or her assessment. However, this did not operate successfully because students frequently used Web-browsers that were configured to prohibit pop-ups. A few complaints were received about the lack of the pop-up because we had informed students that there would be such a pop-up.

During this semester, the software, effectively a beta-version, was frequently fine-tuned. Most of these adjustments to the software improved matters, but not all of them. For example, in the second assignment, most students received a total score that appeared random. However, they could deduce their total score from their

individual question scores. This was fortunate, because there was a delay of some weeks before I was able to send them official notification of their total score for the assessment. There were also two students for whom a number of correct answers had been marked wrong. By probing more deeply into the database it proved possible to determine what mark they should have received. Although OASIS assignments used the same questions that students had seen previously during practice, it made sense to store separately the answers submitted by students to those questions during assessments and during practice. Some software issues arose in achieving this separation, which led to the student responses to one assignment actually being lost. Luckily, the student marks for this assignment were not lost.

Later in this semester, one of my colleagues attempted to run a computer-based test of which half was delivered by OASIS and half by Cecil (see Section 1.2 for an account of Cecil). This was achieved by having half the class do the OASIS part of the test first while the other half did the Cecil part of the test first. Although there were only 45 students concurrently attempting to complete the Cecil part of the test, a part which consisted entirely of multiple-choice questions, the Cecil software completely failed to handle the load. By contrast, the OASIS part of the test ran perfectly. This was a sobering incident for all concerned, given that the OASIS questions were far more complex than the Cecil questions, and that, further, the Cecil software was developed and maintained by several full-time programmers whereas the OASIS software had been produced and maintained by one individual on a part-time basis. Subsequent dialogue with members of the Cecil team produced the reason for the failure: apparently Cecil ran well on some versions of Netscape Navigator 4.7 but very badly on other versions. This response was not particularly reassuring: consequently there was little interest in using Cecil for further assessments, while there was increased conviction that OASIS was a worthwhile product that should be further developed.

7.4.3 The OASIS survey, development and implementation

During this semester I also worked with a colleague from CPD on the production of a survey that would enable me to gain some insight into how students perceived and used OASIS. I did this because I didn't want to ‘reinvent the wheel’, I didn't want the survey to be unduly influenced by my own biases, and I did want to involve someone

in the process who had some expert knowledge in the area. Some of the history of the development of this survey is mentioned in Subsection 5.7.1. The final version of the survey form is reproduced in Appendix one.

My part of the course had taken place in March, while the survey was not finalised till the end of May. Therefore, it was likely that students completing the survey would be thinking more about their experiences in the course and with OASIS in the months of April and May rather than in the month of March. Unfortunately, the parts of the course presented in April and May made less use of OASIS than the part I had presented myself. Further, the OASIS questions for these parts of the course were few in number and often quite repetitive. For these reasons I considered that the survey might show OASIS in a less favourable light than if it had been given earlier. There was also another problem with the timing of the survey. It needed to be given to the students during a lecture but, on the only possible day, all the students had an evening test. Consequently, there was a low completion rate (approximately 40%), and there was also the possibility that the students might not have given sufficient time or attention to the survey. The survey forms were handed out and collected back by the two lecturers who were taking the two lecture streams at the time; one lecturer did use OASIS while the other did not. The survey forms were then analysed by CPD. In the case of the five free-response items, CPD identified all the different student responses for each, along with their frequency. They also returned to me the original survey forms, as completed anonymously by the students, so I could read the student comments myself. About one third of the students who completed the survey were Mechatronics students. Although these students had been fully involved in my part of the course, subsequently they had continued in their own course which did not use OASIS at all. As expected, these students did not respond as affirmatively to the second survey item: “I often use OASIS” as the students taking the full ‘Circuits & Systems’ course did. For this item, 40% of the Mechatronics students agreed or strongly agreed, compared to 81% of the ‘Circuits & systems” students. Apart from this, there was a good match in the responses from the two groups, and the combined responses from the two groups of students are presented in Table 7.4.3.

Table 7.4.3

Student evaluation results for OASIS in year-two classes, semester 1, 2003

Item	SD (%)	D (%)	N (%)	A (%)	SA (%)
OASIS is easy to use	0	1	4	56	35
I often use OASIS	3	8	22	41	26
OASIS helped improve my skill level	0	3	11	45	37
I am more confident about my learning after using OASIS	0	4	20	45	29
The problems provided in OASIS helped me understand the course material better	0	1	15	59	21
OASIS helps me to prepare for the assessments	0	3	9	42	43
I come to campus less often because of OASIS	23	41	17	3	3
I like the instant performance feedback using OASIS	0	1	9	38	48
I prefer to do problems from the text rather than OASIS problems	6	31	48	10	1
It would be a good idea to have OASIS in other courses	1	1	7	47	41

(SD: strongly disagree, D: disagree, N: neutral, A: agree; SA: strongly agree)

In spite of my concerns with the poor timing of the survey, the survey results were definitely encouraging. Most students (91%) agreed or strongly agreed that OASIS was easy to use, while only 1% did not. The instant feedback was appreciated, with 86% agreeing or strongly agreeing with the statement, “I like the instant performance feedback using OASIS”. Furthermore, 82% agreed or strongly agreed with the statement, “OASIS helped improve my skill level”, and 80% agreed or strongly agreed with the statement, “The problems provided in OASIS helped me understand the course material better”, while 85% agreed or strongly agreed with the statement, “OASIS helps me prepare for the assessments”. This latter figure was gratifying, since only 10% of the course assessment was by OASIS, the remaining 90% of assessment being traditional (pen and paper).

It was pleasing to see no indication in the survey that students were coming to campus less often because OASIS provided them with the opportunity to practise questions and sit assignments at home. Only 6% agreed or strongly agreed that they came to campus less often because of OASIS, while 64% disagreed or strongly disagreed. Attendance is an important consideration: keeping students on campus is essential for the University community. The campus is the natural meeting place for students to discuss and solve problems together. They can also easily contact their instructors for help. Keeping students on campus also increases the academic and community spirit and makes it less likely that individual students feel isolated.

In addition to the 10 items set out in the table above, the survey also contained five free-response questions. When asked, "What do you like most about OASIS?" the most frequently mentioned feature was the instant feedback. One student explained it thus: "Instant feedback allows me to quickly identify what I don't know so I can focus my studying." The second most frequently mentioned feature was the fact that students can repeat each question as many times as they like with different numerical values each time. Some noted that this repeatability made OASIS better than a text-book and helped them to really come to grips with the questions. Other common responses to this survey question included the fact that it was easy to use and that practising the OASIS questions led to increased understanding. There was even voiced some appreciation of the motivational power of assessment: one student wrote, "it forced me to learn and do the work more often than I normally would. Regular OASIS assessment is excellent".

When asked "What do you like least about OASIS?" students most frequently noted that the answers are provided without any working or explanation. One student wrote, "It can be a problem to work out a question if you are only given the answer." Two students mentioned being disconnected during assignments. A few students also commented that some of the answers were wrong. This was, unfortunately, true in a few cases. There was also one question (not in my part of the course) that required an algebraic answer but effectively marked A + B correct and B + A incorrect. With time, the percentage of questions for which students are not marked correctly should diminish. One student complained that OASIS provided a numerically different variation of each question on subsequent accesses, whether or not that was actually

wanted: “When you want to retry the question with the same numbers again, OASIS refreshes them each time.” In response to such comments, OASIS does now make it possible to retry the same question with the same numbers if desired.

Overwhelmingly, the most common response to the survey question, “What changes or improvements would you suggest for OASIS?” was a request for worked solutions to the questions, a theme that had also featured strongly in responses to the previous question. One variation was a request for OASIS to “Supply a worked example for each question”. Another student requested “A section that helps you when you are stuck.” A few students took this a step further, with comments like: “Enable us to enter the working and get it marked.” In response to this, a number of students commented, when I spoke to them about other software packages they were using, that they didn’t like filling in intermediate steps when doing problems: they would rather just fill in their answers and, further, they preferred not to have their method constrained to the one implicit in the software. A few students requested statistics so they could see their level of success on each question and compare this to the class as a whole. There were also requests for more variety in the questions, more coverage of the course by the questions, and more questions in general. This theme also emerged in answers to the previous survey question, “What do you least like about OASIS?”

The most common responses to the question, “What benefits do you think students get from using OASIS?” centred on the idea that it provides students with good practice opportunities. It was also seen as providing a large number of questions for student practice and helping students understand. The following is a typical comment: “Increased skill at question-types – understanding of concepts.” One enthusiastic student wrote that the benefits included “Ability to self-assess, focus studying and to know what level they're at. It's great really.” Another noted that OASIS “Fills up our breaks between the lectures in a very constructive manner.” Against the flow of positive responses to this question, two students commented that OASIS offered no benefits beyond those offered by the textbook.

Among the responses to the question: “Have you any other comments to offer about OASIS?”, only one suggestion was made by more than once: the suggestion that OASIS should also be used in other courses. A number of students also wrote encouraging comments such as “It is very, very cool, more statistics would be cooler” and “It is a good learning environment and I think it should be used by other courses.” One student liked the way that OASIS, rather than encourage people to stay home and work in isolation at a computer, actually encouraged group-work: “Using it promotes group involvement often - good to build relationships with others.” Some students also noted that the beta version of OASIS lacked some features which they had appreciated in the prototype: “Bring back the old left side bar, from which you could select any questions from the section you are working in at the time.” This feature, and others missing at that time from the beta version, are now part of OASIS again. A few critical comments were also received, among them: “Although I worked out how to do every question on OASIS, I got a miserable mark in Test One. It doesn't teach you to use principles, rather, you learn a system to solve that specific question.” Most students actually wrote nothing in response to this survey question.

Overall, the feedback from the survey was positive. Students used it frequently, found it easy to use, and considered that it helped their learning. There were criticisms concerning mistakes with some OASIS questions, features that were lacking in the beta version, and the fact that more OASIS questions were desirable, along with OASIS questions in other courses. We were pleased to receive these criticisms because they helped us focus our efforts correctly in our endeavours to improve OASIS. Although a significant number of students stated that they wanted OASIS to provide model answers, or at least worked solutions, to the questions, we were reluctant to provide these, and not just because of the very great effort involved. It was considered that, with model answers, students tend to give up and access the model answers too soon. Further, model answers can reinforce the dangerous notion that there is one best solution method for each question. In contrast to this, I had, more than once, found that students had better ways of solving questions than I had. Further discussion about the provision of model answers and, more generally, the level of feedback that is optimal, may be found in Section 3.5.

7.4.4 Reflections and plans for future action

This semester the students in Circuits & Systems had used a beta version of OASIS which lacked some of the features of the earlier prototype. In particular, a few students stated that they missed the side bar which showed them their average score on each question. Although some useful features were missing from OASIS, its implementation this semester appeared to be most successful. There had been a high number of answer submissions to questions (the median was 169, with a lower quartile of 110 and an upper quartile of 245), indicating a very high usage of OASIS. This indication was backed by our observations and student feedback. The survey conducted at the end of the course indicated that students preferred OASIS questions to textbook questions, found OASIS easy to use, and felt that it improved their understanding. The assignments, delivered by the new server and the new software, ran well without any significant problems even though the overall time period for the assignments was shortened from two days to 12 hours.

Students scored highly on the OASIS assignments, gaining an average mark of 83% for this part of their course work, which compared favourably with an average of 59% for the rest of their course work. Of course, the students should have scored highly on the OASIS assignments, given they comprised questions previously seen in practice. In the final examination the two questions on my part of the course (which was heavily supported by a large number of OASIS questions) returned averages of 71% and 75%. These averages were higher than the averages for the other four questions which ranged from 48% to 68%. Colleagues with knowledge of the course reviewed these two examination questions and judged them to be of a suitable standard, certainly not too easy or too short.

Students also completed a standard course survey: one used each semester with every course in the Engineering Faculty. This survey returned an encouraging number of positive comments concerning OASIS and gave the course a very high rating, just 1% behind the highest-rated year-two course in the Department. In response, the Programme Leader for the Bachelor of Electrical & Electronic Engineering Degree wrote: “This course is an essential Part 2 building-block, but students have often struggled with it in the past. To get a score like this on this kind of course is in my

view a very good result. Congratulations to all concerned!" These sentiments were echoed by the Dean of the Faculty of Engineering.

The trial of OASIS in the course 'Circuits & Systems' in semester one had been most positive. No major problems with the new software had appeared, while any minor problems had been quickly resolved. The new dedicated server had handled all assessment loads well and it was considered that it would do so even for the much larger year-one, semester two class. The students had used OASIS extensively and with a reasonable level of enthusiasm. Overall, students exhibited a high level of learning in the course and OASIS was judged to be one of the key ingredients in achieving this. In the light of the above, the prototype was withdrawn and the new version was made available for all courses that wished to use it. No major changes to the software itself or its implementation were made in going from semester one to semester two. However, the software itself was still being developed and it was expected that small upgrades would be made regularly through the year. I approached the use of OASIS in the second semester with confidence.

7.5 The OASIS Committee

With interest in OASIS increasing, and with the successful trial of the software and the dedicated server, the Head of Department decided to establish an OASIS Committee to better organise and rationalise operations in this area:

> As I announced at our Department meeting I want to set up an OASIS Committee. The role of this group is to overview the present state, use and operation of OASIS, be the coordination point for any new activity involving OASIS, and be the group who will coordinate and advocate for additional resources for OASIS, and oversee the commitment of resources that are provided. (email from HOD, June, 2003)

The Committee was established in the inter-semester break, between semesters one and two. I was appointed as Convenor, along with five other academics, including the previous Head of Department, the Deputy Head of Department (academic) and the Convenor of the Department's Teaching, Learning, Examinations and Assessment Committee. At our first meeting, we decided on our terms of reference and these were subsequently approved by the Head of Department:

(1) to monitor and review the present state, use and operation of OASIS

(2) to be the coordination point for any new requests and activities involving OASIS
(3) to set priorities and dates for future activities and developments
(4) to advocate for continued and additional resources for OASIS
(5) to oversee the commitment of resources that are provided.

To clarify the above, it was agreed that activities involving OASIS could include:

(a) writing and encoding questions
(b) writing software for OASIS to enhance its functionality for students, staff, other departments, and researchers
(c) carrying out research to confirm its positive impact on student learning
(d) establishing guidelines or rules for staff who use OASIS and/or write questions for OASIS
(e) implementing training sessions for instructors who use it and students who encode problems
(f) producing documentation for users, developers, question-writers, question-encoders,…
(g) producing and implementing a disaster recovery plan
(h) promoting OASIS.

The membership of the OASIS Committee has not changed from its inception: the Committee still has the same six members and I am still the Convenor. The above terms of reference and list of activities still relate well to the business of the Committee, although one or two additional areas of concern have arisen since the Committee's inception; for example, there has been some debate about whether the University should attempt to market software packages such as Cecil and OASIS, or whether OASIS should be provided as an open source software package.

7.6 Large-class assessments: Semester two, 2003

7.6.1 Introduction: Changes from 2002

This section describes the implementation of OASIS with the year-one class ‘Electrical Engineering Systems’ in the second semester of 2003. Some background information about this course may be found in Subsection 6.4.1. Compared to 2002, the number of students had risen slightly, to 560, and the cut-off level for admission to the course (based on the national external UEBS examinations, sat by students at the end of their final high-school year) had been set at a slightly higher level.

There were, however, some significant changes from 2002 in the course lecturers. The Deputy Vice-Chancellor (research) no longer had a role in the course; instead I

presented the 'Introduction to electric circuits' section he had presented in 2002. I also taught the 'Power systems & electrical machines' section that I had taught in 2002. Altogether, I lectured about 45% of the course. The only other lecturer retained from 2002 to teach the course again in 2003 gave just five lectures. The Convenor of the Department's Teaching, Learning, Examinations and Assessment Committee taught two sections of the course, 'Logic systems' and 'Operational amplifiers & transistors', about 45% of the course. The content of these two sections was also changed somewhat. In my parts of the course there were no large changes but, in response to students taking longer than expected to grasp some concepts, I removed some material to allow more time to be devoted to the key concepts.

There were also significant changes to the course assessment for 2003. I was keen to include some OASIS assignments, and the trial of OASIS in the first semester indicated that the system could handle an assignment for the very large number of students involved. The final grade in this course was derived from the final examination (70%), two traditional written tests (each 10%), and two OASIS assignments (each 5%). In fact, the original intention had been to have three OASIS assignments worth a total of 10%. However, sufficient questions were not ready in time to cover enough of the course to enable all three assignments. Consequently, the number of the assignments was reduced to two; each assignment covering one of my two parts of the course. Previously 10% of the course had been allocated to a formal laboratory report. While the same laboratory work was still required in 2003, no formal report was necessary. Instead, the laboratory work contributed to students' final grades through a compulsory question in the final examination. Altogether there were six questions in the final examination, one for each of the five sections in the course, and one covering the laboratory work.

7.6.2 The OASIS assignments

As outlined above, originally three OASIS assignments had been planned. The first of these had been scheduled for the third week of the twelve-week semester. I was keen to have an assessment early to motivate the students to settle into effective study habits from the start of the course. Further, an OASIS assignment was ideal because it provided instant feedback, giving students, early in the semester, a realistic assessment of how their achievement matched the required standard. However, this

assignment did not eventuate because the required questions were not produced in time. Consequently, the first assessment was a traditional written test in week five of the semester which was not marked and returned to the students until the second half of the semester. The two OASIS assignments that did take place, based on my two sections of the course, were held in weeks seven and 11 of the semester. Unfortunately, this timing was too late to provide early motivation or early feedback. However, the OASIS assignment held in week 11 of the course did enable students to receive some feedback on the latter part of the course prior to the final examination. This would not have been possible with a traditional test because the marking would have taken too long. Typically, in this course, students waited about three weeks for a traditional test to be marked and returned.

I introduced the students to OASIS in week four of the semester, soon after I started lecturing. At this time there were significant demands on their time. Perhaps this was the reason that there was not a great deal of OASIS activity until about one week before the assignment. In the week leading up to the assignment I received e-mails from about 30 students seeking help with OASIS questions. Answering these emails became quick once I had drafted a set of responses that could be copied and pasted to form the basis of replies. A number of students also spoke to me before or after lectures, or during office hours, about the OASIS questions. Many students also sought help during their tutorials. Some questions were clearly giving students particular difficulties and these were sometimes addressed in lectures.

There was a large increase in the number of answers submitted to OASIS questions in the last few days before the assignment. The assignment itself was scheduled to start at 8 am on a Monday. By the previous Friday, there had only been 10 000 answers submitted. By midday Sunday, this figure had increased to 17 000, and by the start of the assignment the figure had increased further to 27 000. This last figure represents an average of approximately 50 answers submitted per student in the class. The figures also show that, of these 50 submissions, a good 30 of them actually took place in the weekend prior to the Monday morning assignment. These figures do provide definitive evidence for procrastination on the part of students. While instructors often suspect that their students leave tasks till the last possible moment, seldom do they receive such clear confirmation that this is true.

The assignment itself ran reasonably smoothly with no major problems, although there were a few minor problems. For example, two students received incorrect total scores, while a few students answered only the first question and then submitted their assignments. This required resetting their assignments to give them a second chance to complete it. To avoid this in the future, perhaps a warning should appear when students attempt to submit their assignments, or perhaps a practice assignment should be provided prior to the first real assignment. During the course of the assignment, the server and network handled the load adequately with no slowing of the system. Fortunately, this was helped by the fact that the students spread their assignment times fairly well through the allocated twelve-hour period. The dedicated server and the improved OASIS software also helped make it possible for the assignment to be done by such a large group of students over the twelve-hour period: testing revealed that the new OASIS was three times faster than the prototype.

The second OASIS assignment took place four weeks after the first one. Again, there was a remarkable increase in the number of answers submitted to OASIS questions just prior to the assignment, as indicated in Figure 7.6.2.

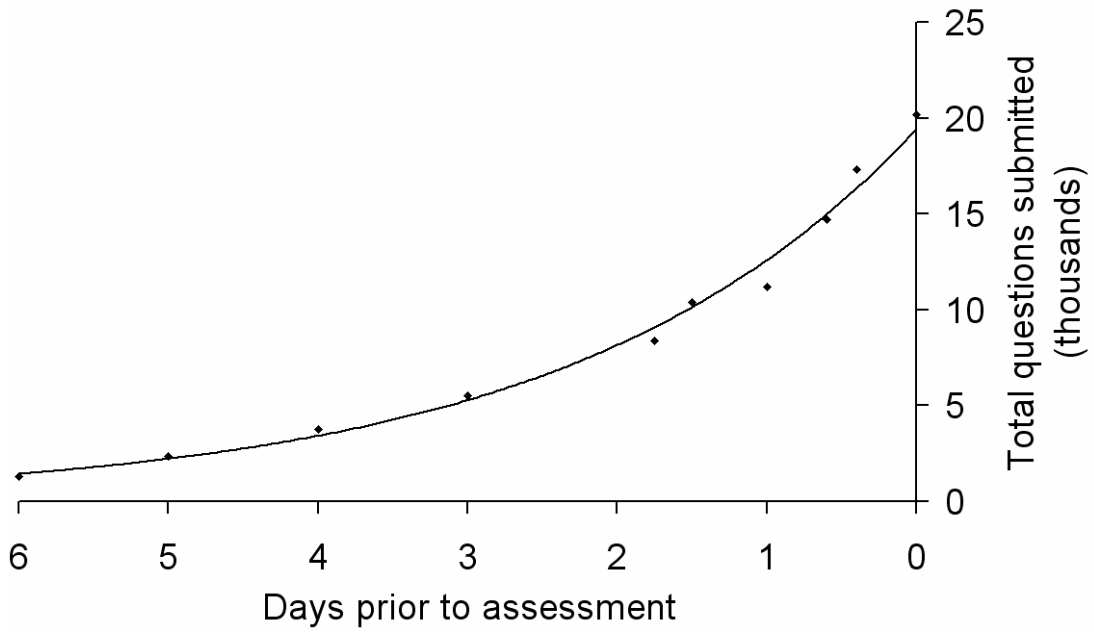


Figure 7.6.2: Year-one student activity prior to the second assignment, 2003.

This assignment was the second sat by the students, and, as hoped, none submitted their assignments before they intended. In fact, the only significant problem with the assignment was one that had nothing to do with the students, the instructors, or the software: the University Internet connection went down for about one hour on the evening of the assignment. Fortunately, this only affected 10 students and a re-sit assignment was run for them.

In general, students scored highly in both assignments. In the first assignment the average mark was 82%, in the second, 92%. In the first assignment, 200 students had scored full marks, with a further 87 students scoring 15 out of 16. The corresponding figures for the second assignment were even higher: 310 and 100. The second assignment marks were expected to be higher because there were fewer questions for the students to work through in this section, there were no very demanding questions (as there had been in the first section), and the students were more familiar with OASIS assignments. Of course, it may also have been that, by the time of the second assignment, more students had worked out how to 'beat the system'.

7.6.3 The OASIS survey

Ten days after the second OASIS assignment, the students were asked to complete a survey on OASIS. This timing put the survey on the last day of the course, just eight days before the final examination. In total, 333 students completed the survey forms. The survey was the one used with year-two students in semester one, as outlined in Subsection 7.4.3. Again the survey forms were analysed by CPD and then returned to me so that I could read the comments in the free-response section. Table 7.6.3 summarises the data from the first page of the survey form.

Although the students were asked to complete the survey on the last day of the course, a day when they might well have had other things on their minds, the survey results were positive. Most students (88%) agreed or strongly agreed that OASIS was easy to use, while only 2% did not. These figures were very similar to those returned by the semester-one, year-two class (91% and 1% respectively). The instant feedback was appreciated, with 80% agreeing or strongly agreeing with the statement "I like the instant performance feedback using OASIS". The corresponding year-two figure had been somewhat higher at 86%.

Table 7.6.3

Student evaluation results for OASIS in year-one classes, semester 2, 2003

Item	SD (%)	D (%)	N (%)	A (%)	SA (%)
OASIS is easy to use	1	1	5	45	43
I often use OASIS	8	20	41	21	6
OASIS helped improve my skill level	2	7	18	53	18
I am more confident about my learning after using OASIS	2	8	28	47	12
The problems provided in OASIS helped me understand the course material better	0	7	14	55	19
OASIS helps me to prepare for the assessments	2	5	16	43	31
I come to campus less often because of OASIS	27	39	17	3	3
I like the instant performance feedback using OASIS	1	2	11	35	45
I prefer to do problems from the text rather than OASIS problems	11	31	34	10	6
It would be a good idea to have OASIS in other courses	3	4	12	43	33

(SD: strongly disagree, D: disagree, N: neutral, A: agree; SA: strongly agree)

71% agreed or strongly agreed with the statement "OASIS helped improve my skill level". This result was down on the year-two figure of 84%. Continuing the trend, 74% agreed or strongly agreed with the statement "The problems provided in OASIS helped me understand the course material better", whereas the year-two figure had been higher at 80%. In general, the year-one students, while positive about OASIS, were somewhat less positive than the year-two students. There are two plausible reasons for this that do not relate directly to OASIS or its implementation. First, in the year-one course, OASIS had not been introduced until week four of the semester. By that time, students were busy meeting many deadlines. In contrast, the year-two students had met OASIS in the first week of the semester, a time when there were no conflicting demands on their time. Second, the year-two students were all majoring in electrical engineering and so were more motivated to understand the material. By contrast, most of the year-one students were not majoring in electrical engineering

and so were likely to be less motivated to learn the material in a deep way. In fact, a number of students had not wished to study electrical engineering at all in year one and so tended to be negative about anything related to electrical engineering. In spite of these considerations, 74% of the students thought that OASIS had helped them prepare for assessments (90% of which were PPA), and 76% of the students thought that it would be a good idea to have OASIS in other courses. Following the year-two pattern, the textbook problems were also less popular than the OASIS problems, with 16% agreeing or strongly agreeing, and 42% disagreeing or strongly disagreeing, with the statement "I prefer to do problems from the text rather than OASIS problems".

As before, the survey also contained five free-response questions. The year-one response to these questions was not as full or as considered as the year-two response. Approximately half of the students wrote nothing at all in response to these questions. Of those that did respond, some wrote one-word responses such as "everything" or "nothing". When asked, "What do you like most about OASIS?" the most frequently mentioned feature, and by a very large margin, was the instant feedback. This was also the most frequently mentioned feature in the year-two survey. The second most frequent comment was that OASIS was easy to use. The third most frequently mentioned feature was the fact that students can repeat each question as often as they like with different numerical values each time. One explained the benefit of this as follows: "I can repeat a certain type of question many times with different values so I can be sure I totally understand how to do the question by practising it many times". Several others wrote similar comments. This theme had also emerged in the year-two responses to the survey. Several students noted the fact that OASIS can be used from home, even for assignments.

When asked, "What do you like least about OASIS?" the year-one students, following the year-two pattern, most frequently noted that the answers are provided without any working or explanation. One student commented: "Only the answer matters. There's no way for me to know where I have gone wrong in the process if I have incorrect answer". The second most frequent theme in the year-one responses was the fact that people can cheat on assignments. This theme had not been mentioned in the year-two survey at all. One student outlined the problem as follows:

“People cheat way too much in the assessments. In the computer labs groups of people do the test together and others distribute Excel spreadsheets containing solution methods”. Several students also mentioned problems related to computers: “It’s on the computer, and computers are prone to stuff up”, “Dependent on a reliable Internet connection”. There were also requests for more OASIS questions and for more areas of the course to be covered by OASIS. Some comments were reminders that there are two sides to every story, for example: “It is really hard to use and also really complicated”, and “You can just learn how to get the answer for a specific question and kid yourself you understand the concept”. The latter comment does highlight an important issue. It is certainly possible for a student to solve a problem in a mechanical way and still get the right answer. Quite likely, sadly there are many students who do pass courses this way. Course assessments should be constructed so that, overall, sufficient novel questions are used so that rote-learners are unlikely to succeed. The OASIS assignments, by contrast, did rely on previously seen questions. Students could learn or write out the methods for these and pass the assignment. However, OASIS did contribute only 10% to the final course grade. The other 90% did include more searching, previously unseen, material. The rationale for the OASIS assessments was that they would, it was hoped, motivate the students to master the OASIS questions and, in doing so, the students would lift their skill levels considerably. It was always acknowledged that OASIS, while it could make an important contribution, would never be a complete learning or assessment package.

As it had been for the year-two survey, the most common year-one response to the question “What changes or improvements would you suggest for OASIS?” was a request for worked solutions to the questions. This response was in agreement with the responses to the previous question. One student suggested that OASIS should “show how you get the answer after a few failed attempts”. Other students suggested that OASIS should have a hints section, explanations of how to get the answers, and links to the relevant theory. More practice questions and more variety of topics were each requested almost as frequently as worked solutions: “more varied problems, i.e. from different topics such as transistors and more difficult test/exam question section”, “it’s applicable to many other courses and would be great as a tool for learning in MM1 [year-one mathematics] and Engen121 [year-one mechanics]”. A few students asked specifically for some more challenging questions.

By an overwhelming margin, and mirroring the year-two response, the most common replies to the question "What benefits do you think students get from using OASIS?" centred on the idea that it provides students with good practice opportunities. Other frequently-mentioned benefits were: improved learning, understanding of electrical circuits and problem-solving knowledge: "learn lots that they wouldn't have", "understand concepts", "improved learning", "problem-solving skills", "improved skills", "increased skill knowledge", "can practise the same sort of questions until you understand", "an ability to ensure you understand the problems before assessment", "can test their knowledge and they can make sure they understand the content correctly", "identify problem areas", "they instantly get told where gaps in their learning are", "learn the basic understanding of the topic", "being able to go over sample problems over and over until the finer points of the course material is firmly embedded in a practical way". A few students mentioned as a benefit the fact that OASIS was "forcing students to practise". The instant "feedback on learning" was also mentioned. A few students considered that OASIS did not provide a great benefit to students "because people just make formulas with limited understanding". Certainly, while some students had practised frequently, lifting their skill levels, others had done minimal practice, taking a formulaic approach or relying on the spreadsheets of others.

The most common response to the question: "Have you any other comments to offer about OASIS?" was an encouraging statement to the effect that OASIS was good. As for the year-two survey, another popular response to the question was the suggestion that OASIS should be used in other courses, or other parts of the course: "I will love Electrical if and only if you can provide more practice for every topic learnt in the course so I can understand this subject better". There were few ideas expressed here that had not been already expressed in responses to previous questions. One new idea involved introducing multiple-choice questions to test concepts: "For OASIS, questions involving calculations may not be appropriate. I recommend the questions to be multiple-choice which test our concepts."

Overall, the survey feedback was positive. In general, students used OASIS often, found it easy to use, and stated that it helped advance their learning. While the survey

was not as positive as the year-two survey, the difference was judged to be not greatly significant. It was considered that it could be explained in terms of the facts that the year-one students had been introduced to OASIS later in the course and that they were less well-disposed towards Electrical Engineering than the year-two students, for all of whom it was their major. Following the year-two pattern, many year-one students wanted worked solutions or hints for the questions. However, we did not wish to provide these, for the reasons outlined in Subsection 7.4.3.

7.6.4 Reflections

In 2002, even without the motivation of OASIS assignments, the majority of the year-one students had used OASIS extensively, with the average student submitting about 80 answers to OASIS questions, questions that covered 40% of the final examination. In 2003, the same questions generated an average of about 130 submissions, with 110 of these prior to the second OASIS assignment (18 days before the final examination). The assignments would clearly seem to be the main factor for the increase in answers submitted. In 2003 the OASIS questions covered only 33% of the examination owing to a format revision and so the motivating factor of the final examination was actually reduced.

The survey conducted at the end of the course indicated that, in general, students found OASIS easy to use and considered that it lifted their skill level and understanding. A preference was expressed for OASIS rather than textbook questions. While the survey results were certainly favourable, they were somewhat less so than the year-two survey results. It was judged that this difference could be largely explained in terms of the students involved: most of the year-one students were not majoring in Electrical Engineering whereas all the year-two students were.

The only significant issue to appear in the year-one survey that had not appeared in the year-two survey was that of ‘cheating’ in the assignments. My colleagues and I were not greatly concerned about the sort of collaboration in which all students contributed to a reasonable extent. We believed that in this situation all students would be learning and improving their skills. Nor were we concerned about the situation in which a student had constructed a spreadsheet in advance and then subsequently used it for the OASIS assignment. We considered that, to do this, a

student would need a good understanding of all the practice questions. However, we were concerned about students who relied on spreadsheets produced by others: such students could well decide to do very little OASIS practice, knowing that they could pass the assignments simply by using the spreadsheet. This practice would certainly defeat the purpose of our implementation of OASIS. It would be important to limit, as much as possible, the effectiveness of the trade in spreadsheets. This trade was much more prevalent in the year-one class than in the year-two class. Possible reasons for the greater level of cheating at year one included the lower maturity and the lower interest in Electrical Engineering, the fact that the assignments had taken place at a busier time of year, and the fact that the assignments contributed more to the final grade, 5% each. In contrast, at year-two level, there had been four assignments worth a total of 10% of the final grade. Appropriate ways of minimising the trade in spreadsheets are discussed in Section 7.9.

The assignments, as in the first semester, ran well without significant problems, even though the number of students involved had risen dramatically to 550, with no increase in the time allocated for assignment completion. Students gained an average mark of 87% for the OASIS assignments. This high mark was to be expected, since the students had had the opportunity for prior practice of the questions. By comparison the average mark for the two tests was 62% and, for the final examination, 58%. In this examination the two questions on my part of the course, supported by OASIS questions, returned averages of 54% and 68%. The other four questions averaged 56%. The pass-rates for my two questions were 61% and 84%. While these figures were comparable to those of 2002 and maintained the large improvement over 2001, I believed there was clear room for further improvement in the section of the course ('Introduction to electric circuits') that had returned the poorer statistics. 2003 was the first time I had taught all this section, and I judged that it contained too much content, encouraging some students to learn the material superficially. For 2004, my aim was to reduce the content and aim for deeper student understanding of what remained. I judged that 'Power systems & electrical machines', the other section of the course I had taught, and had previously taught in 2002, had gone well. Here I had trimmed the content. One third of the relevant examination question had required written explanations while the rest involved calculations. This latter part, supported by the OASIS questions, had been done

extremely well by the vast majority of students. By contrast, a number of students had struggled with the written part of the question, but no more than expected, given that many of the students did not have a strong command of the English language.

7.7 The first OASIS student interviews

Student interviews were conducted during the second semester of 2003 and augmented the data from the student surveys. Unless indicated otherwise, all indented text in this section consists of quotes from interviewed students. Students may also be denoted by S. The interviewer is denoted by CS. Seven interviews were conducted with students from the first-semester year-two 'Circuits & Systems' class, the interviews taking place after their OASIS survey was conducted and analysed, providing an opportunity to explore issues highlighted in that survey. The background for the interviews has already been covered in Subsection 5.7.2. In particular, I noted there that, although I emailed a good range of students, in terms of both their academic achievement and extent of OASIS usage, I received a relatively low response rate, even after follow-up emails. This led to a different approach with the second-semester, year-one class, namely an emailed request for volunteers. This request produced another seven interviews. Thus, altogether 14 interviews were conducted in semester two, with a typical interview lasting 15 minutes and producing a transcript of around 2000 words: four or five pages. The software package N6 was used to help analyse these transcripts. Although the group of students interviewed was not representative of the whole class, the exercise was a valuable one because it did yield a number of useful insights into how students perceived and actually used OASIS. The insights went well beyond the 14 students interviewed since these students frequently commented on the practices of their peers. The interviews also enabled some survey issues to be explored in more depth, reinforcing and expanding on information received from the survey. For example, the survey had indicated that, in general, students preferred to use OASIS rather than the textbook. The interviews revealed more about the reasons for this preference. Some of these reasons centered on the convenience of OASIS and the inconvenience of carrying the textbook around, or borrowing it from the library on short-term loan:

> It's helped me out a lot. Yes, I have used it a lot. I just find it more convenient than lugging around a two-kilo textbook everywhere.

> It's also easy to go down to the computer because you are on the computer more often than you go into the library. So, definitely more time spent using OASIS.
>
> You can just get to it any time you want to and you don't have to worry about going to the library and making bookings for the library book and when you want it, it's there - that's the good thing about it.

Even though I had set problems from the textbook for the students, and they did have access to answers for all of these questions, students still felt that the way in which OASIS marked their answers was quicker and easier than the methods provided by the textbook:

> If I got it wrong, very often a text book doesn't have an answer or the answer is very hard to find or we could go on the Web or receive it from somebody.
>
> The text book doesn't even have the answers. Most of them. Unless you use the CD ROM. A lot of computers around here don't actually have the CD ROM drive.
>
> I spend more time on OASIS, more than the text book because I found, with OASIS, you get the answers faster.
>
> I prefer OASIS all the way. It's much faster.

However, not every student voiced a preference for OASIS questions. One student stated he preferred textbook questions, possibly because he had a preference for questions on paper as opposed to questions on a computer screen:

> Personally I printed out every question of OASIS. In a way, I suppose it could be possible to give printouts of the OASIS questions to the students. That could also be an alternative, but it doesn't have to be done. Personally I prefer textbook questions.

When asked to suggest improvements for OASIS, a number of students did indicate that they would like to see model answers, or at least some sort of assistance or references provided, for the questions:

> And one more thing - you can have a model answer for the questions. That will be good, because sometimes you are really, really stuck and you have no idea where to go and sometimes it's during the break that you are doing it and for people who live on the North Shore, it's really hard for them to come all the way down, just to ask one small little question.
>
> But if I could print out things with answers, model answers, and take that away as another, that could be one extra way I could use it. One other thing would be to perhaps reference Hambley or the course-book page numbers, along there, or if there was another layer of hints, like formulae if you need it.

However, not all students were in favour of solutions. One voiced a preference for minimal feedback:

> I like that it tells me it's wrong and I like that it doesn't give me a solution, because the tendency is when you see a solution, you tend to go on to the next one.

In general, students felt that the ability to repeat OASIS questions with different numerical values was a significant benefit:

> Yes, getting different numbers is really good. It means you have to do that method again. It's really tedious when you are doing it but later on you realize that because you have done it 15 times, you're actually pretty proficient at it.
>
> What I usually did was - I did the whole thing, if I got it right, I moved on to the next one otherwise I'd do it all again with different numbers, just to make sure.
>
> You've got to know the theory behind the problem, to answer the question and then, if somebody else has to explain that to you, then that's a lot better than giving the answer isn't it?

Some students reported that they would repeat questions five or six times. One student, quoted above, mentioned some questions 15 times, while another aimed for a particular score on each question, as the following reveals:

> CS: If you weren't too sure how to do a particular problem with OASIS, you could do it more than once but with different numbers. Was that of any help to you?
>
> S: Yes, if you want to make sure that you understand a problem, aim for 90% right. That is 10 times you do the problem and 9 times you get it right. You can try that on OASIS, and achieve the goal.
>
> CS: So did you actually do that with some problems? If you got them wrong once, would you say, now I've got to get it right 9 times?
>
> S: Yes.

There was no clear location preference stated for using OASIS: some students preferred to work at University, others at home. It was simply a matter of the most convenient location. Almost all students had access to the Internet at home and OASIS performance seemed adequate there even though the University Internet connection was faster than home broadband and much faster than home dial-up.

In general, students preferred to work on their own when using OASIS. However, group work was often resorted to when difficulties where encountered:

> I have always worked on my own. I have found that if I have friends or other people around, I cannot concentrate, so I prefer to go in a corner where nobody else is.

> At first I'll just try them by myself, try to answer all questions and see what I can do and what I can't do. Basically the questions that I find difficult, I will discuss with my friends.
>
> Work on my own, unless somebody is stuck, then I'll work with them. Because I don't see the point of working with other people, especially when there is a limited time. OASIS problems have limited time. But I usually do them on my own, at home. It's much more effective for me.
>
> And I thought the difficulty of them [the questions] was good too. They started quite easy and then later became quite challenging. That was good. You had one or two questions in there where everyone would get stuck and there would be a big discussion about it, because everyone is doing the same things, in practice... With OASIS, because everyone is doing the same questions, you can talk to others and get it sorted out.

One student, by contrast, did voice a preference for group-work and described how the group went about it:

> CS: Would you usually work on your own or would you work with some other people on the OASIS questions?
>
> S: Usually work with other people and try and get through them, step by step. One by one, go through the practice questions and that makes it a lot easier because you've got five people figuring out how to do it.
>
> CS: So that means you would be sitting at the same computer, at the same time, working at exactly the same problem, or…?
>
> S: No, it's too difficult to try to do it all on one computer. What we usually did was have four or five people and just logon on different computers and do individual questions, but we can work out how to do them together.

While there was general enthusiasm for OASIS as a learning tool, some students did voice disquiet about its appropriateness for assessment. There was a feeling that some students had been able to beat the system by relying on the efforts of others:

> Some people are doing too well on the assignments: most of them are probably working together.
>
> When I did my OASIS assignment, I saw other people actually grouping, gathering, 3 or 5 people. They're doing everybody's assignment at the same time.
>
> Also another thing I observed quite lot for the OASIS test were students collaborating quite a bit. That I think is quite a problem. I've even seen students, I've heard rumours, of all these spreadsheets going around. That's what I've heard.
>
> I have seen the spreadsheets being passed around the room. I have seen group assignments. I think that part is the weakest link in the OASIS… Myself, I write down every single item and I can basically work around paper numbers

on the calculator. It's not much of a test. I actually get far more out of the practice than the test.

In general, students considered that allocating 10% of the final grade to OASIS assignments was appropriate. They considered that, because the assignments were unsupervised, the allocation should not be higher than 10%. The figure of 10% was accepted as providing an incentive for students to practise the OASIS questions and learn the material:

CS: Do you think 10% for OASIS is about right or do you think that number should be different?

S: I think that's okay at the moment, because you don't really know with OASIS, whether each person is actually doing it or some other person or many people doing it together.... Don't get rid of it. I think that 10% is okay.

Even if there are some people cheating it isn't too bad, in the sense that, if they are going to cheat to get through, they are going to get caught at other barriers, like tests and exams. That's not going to get them through.

CS: About 10% of the course marks come from OASIS. Now do you think that's about right or do you think that's too high?

S: I think that's about right because the assessments themselves probably don't really show anything because they are open book. It's the fact that you have to practise for them, to get the marks and it's the practising that helps. It's the actual practice that makes you good at things.

I think that 10% is quite good. I mean the 10% is not too hard to get 8 or 9 out of 10. So if you make it smaller then people might just not do it.

In the interviews, I explored whether students would prefer a supervised test or an unsupervised assignment. There was no clear preference voiced for either. Some students preferred the idea of a supervised test because it would eliminate collaboration and cheating during assignments. Others preferred the status quo because they found tests highly stressful. As suggested above, if more than 10% of the course marks had been assigned to OASIS assessments, there would have been a noticeable preference for supervised tests:

Doesn't matter whether it is assignment or test, I think people still use it. They find it useful.

I can't see any problem with that [OASIS tests]. That would be good. Though I wouldn't ditch the assignments for that, I'd still have the assignments. That's a great way to learn.

I think the idea of it being in a more test environment is a good idea. Obviously you've got to find five hundred odd computers.

> I don't recommend that [supervised assessments] at all because that will be more like a test. So there won't be an assignment and then the whole concept of an assignment will be gone. There's nothing really that can be done about it - people work together, as long as they don't get someone else to do it for them, it's ok.

7.8 The first OASIS staff interviews

To increase triangulation, the decision was made to obtain an independent evaluation of OASIS conducted by someone recognised as knowledgeable in the evaluation of CAL. As part of this evaluation, and as outlined in Subsection 5.7.6, a colleague from CPD interviewed all staff members who were currently or had ever been involved in some significant way with OASIS. This evaluation increased triangulation and therefore the rigour of the research in a number of ways: first, another researcher was involved; second, she used her own methods; third, she interpreted the data she collected; and fourth, she formally interviewed a group - the academic staff - that I had not formally interviewed. These interviews took place in the inter-semester break in 2003. Based on the findings from these interviews, a report entitled 'OASIS Impact Evaluation' was produced. It was most positive:

> OASIS has already proved to be a highly efficient and effective system for delivery of formative assessment tasks and feedback to students in ways that would not otherwise be possible because of the large class sizes involved.... in relation to 'state of the art' computer-based assessment systems, OASIS rates among the best available and is more sophisticated functionally than many commercial products. This is true of the educational as well as the technical design.

The report also noted that OASIS tended to encourage student collaboration and that one consequence of this was decreased demands on staff time:

> Instead of having to deal repeatedly with the same basic questions from individual students, staff time can now be devoted to helping small groups of students with specific problems after they have worked through them in OASIS, received feedback on errors and discussed them with peers.

Also raised was the possibility that "standards could be set higher because of the learning gains achieved through integrated use of the system".

Later, the academic who had conducted the staff interviews made available to me her interview notes which went into much more detail and raised some questions for future investigation. These notes began by explaining the aims of her evaluation:

> The principal aim of the evaluation was to investigate system functionality and performance. A secondary aim was to provide a summary report of the overall impact of the use of OASIS computer-based assessment on both staff and students involved in the courses.

With respect to the principal aim, OASIS was seen as setting a high standard, both in terms of its ability to produce numerically different versions of the same questions, and in terms of its ability to run satisfactorily on a wide range of browsers:

> From a technical perspective, OASIS offers complex functionality that allows random generation of questions so students can practise on similar questions with different values. In this and some other respects, it is considered a product superior to many proprietary brands.
>
> OASIS works on all popular browser versions, achievement of this is a technically complex issue.

However, the point was also made that the sorts of questions that could currently be offered by OASIS were someone limited. Therefore, it should be seen as a partial rather than a complete solution to assessment and learning requirements. In particular, there was the suggestion that it may be better suited to year-one and year-two courses rather than more advanced courses:

> OASIS is considered useful as part of an educational context that includes other types of assessments and tests, e.g. written tests to assess analytical skills.
>
> The version of OASIS in use at the time of the interviews was considered ideal for years 1 and 2 whereas year 3 requires 'higher order' skills that may not be well served by the forms of assessment possible in this environment. Future versions are expected to offer greater flexibility.

Of concern was the fact that, while staff and students found OASIS easy to use, the difficulty of uploading questions to the OASIS database remained. In fact, most staff did not perform this task; it was instead carried out by graduate students:

> OASIS has proved uniformly easy to use for staff and students alike. However, most staff are not able to write and input questions so the idea of employing graduate students on contract is suggested.

At this time, graduate students were in fact paid a flat rate of $50 per question. Consequently, producing a set of OASIS questions for a course that had not previously used OASIS could easily have entailed a cost of a few thousand dollars.

In addressing the issue of the impact of OASIS on student learning, it was noted that:

> All staff interviewed report that the overall student response to OASIS is positive. This is backed up by observation and system log data that shows they use the system extensively.
>
> The instant feedback feature is a valuable aid to student learning.
>
> Relevance of OASIS questions to coursework and assessment is obvious to students.

The advent of OASIS enabled courses to have regular assignments. The consequence of this was a more regular work pattern on the part of the students. Without automated marking, such regular assessments would have been impossible and students would have tended to put off serious learning till much later in their courses:

> Learning and reinforcement is now spread out over the semester rather than being covered once then crammed in before exams.

There was also some evidence that OASIS had modified the way students approached their learning. In particular, it was conjectured that students had managed their own learning better and had become less dependent on the instructor. Student collaboration also appeared to have increased:

> Students may be teaching themselves to work through problems on OASIS. This could lead to better autonomy and self-management as well as catering to a wider variety of learning styles than other approaches.
>
> The type of questions that students ask the staff have changed. Previously, they were simple "how do I do this?" type questions from individual students. Now they come in groups with more specifically focused questions, i.e. what process or method are they failing on. The assumption is that they first try to solve problems with OASIS and with peers, then as a last resort come to the lecturer for help.
>
> Students seem to be encouraged to collaborate more while using OASIS, this is either face to face or with chat sessions running simultaneously.

There was a belief in the Department that OASIS was raising the standard of learning and that correspondingly higher standards could therefore be set. Some staff judged that this had already happened without those involved being particularly aware of it:

> Now students are using OASIS to support learning, it may be possible for the goal posts to be raised, i.e. for overall higher standards to be set and achieved within the allotted timeframes. It was suggested by some staff that this has already happened. The implications of this require due consideration, i.e. will UoA graduates be more attractive to employers, will degrees be seen as exceptional quality and would this have any impact on recruitment and selection of students?

It was also observed that even staff initially sceptical or disinterested in OASIS now accepted that it was of benefit to both staff and students. However, though there was

general agreement that students had to learn methods rather than memorise numbers, there was some debate as to how deep the learning promoted by OASIS really was:

> Some staff voiced concerns that they did not know if OASIS was promoting surface or rote learning that would not be retained in long-term memory or transferred to high performance ability in later years of degree programs. Others believed that the questions were well structured so this would not be a problem.

The notes also suggested a few ways in which OASIS and its implementation could be improved. First, there were some software and question errors that should be corrected: however, in this regard it appeared that students were generally prepared to cope with minor difficulties because they considered that the system was useful. Second, further development was required so that OASIS could cover a larger variety of questions and offer more marking options. Third, more questions were needed so that OASIS could better cover courses in which it was already used and, further, be implemented in courses where it was not currently used. Fourth, further work needed to be done to measure the impact on student learning in the long term to "ensure transfer and competence at higher levels are being adequately supported".

Finally, the notes concluded by listing questions and issues for further research. Some of these have been mentioned already. Further issues are raised in the following questions:

> What is the best way to integrate the benefits of OASIS into the overall course 'landscape'? I.e. which parts does OASIS reach and which parts does it not?
>
> What characteristics of 'good questions' will address concerns about rote learning?
>
> How much student effort goes into finding 'workaround' strategies rather than working on problems and can this be shifted to good effect?
>
> Are some students NOT using OASIS and does this have an impact on their results?

Altogether, the report provided a good deal of positive feedback and a strong validation of the work that had been done by the academic staff in producing and implementing OASIS. It also highlighted areas for improvement and issues for further consideration and research. For example, to what extent could good design of OASIS questions minimize rote learning and maximize understanding? And, what sort of questions and implementation policies would encourage as many students as possible to actively engage with OASIS?

7.9 Reflections and plans for future action

After a successful trial with one class in the first semester of 2003, the new version of OASIS completely replaced the prototype version in the second semester. OASIS was well used by the students in courses where it was made available, and it met with a good level of approval from both students and staff. Data collected by OASIS itself attested to the high level of student use, while both surveys and interviews were used to establish the student reaction to OASIS. An independent study conducted by an academic from CPD and documented in the report 'OASIS Impact Evaluation' concluded that OASIS was benefiting the learning of students as well as freeing up staff time. By the end of the first semester there was considerable evidence to suggest that the department should continue to support and develop OASIS. To this end, a committee was established, with myself as the convenor.

In the report 'OASIS Impact Evaluation' it was noted that "OASIS... is more sophisticated functionally than many commercial products". While this may well have been true, the reality was that the software was in a state of constant change and improvement. This was a result of ongoing reflection and action through the year. By way of illustration, some of the most significant features added to OASIS during the second semester of 2003 are now outlined. First, the prototype version had the marking tolerance fixed at 1%. Such a small tolerance was inappropriate for some questions, such as those requiring students to read data off graphs. Consequently, the software was revised to allow questions to be marked to a specifiable tolerance, by default 1%. Second, while variable values could be placed on diagrams, they could not be placed within a body of text. Consequently, questions often needed to be written in a somewhat stilted form:

> An ideal voltage source of V volts is connected to a resistance of R Ω. Find the current through the resistance, given the following values:
> $V = 13.5$ V
> $R = 10$ kΩ.

This sort of wording was considered most unnatural. The use of symbols suggested to students that the questions were formulaic; further, the specific symbols used provided clues to appropriate formulae. With additional programming, it became possible to place variables in text. The above example now became:

> An ideal voltage source of 13.5 volts is connected to a resistance of 10 kΩ. Find the current through the resistance.

Here it should be appreciated that the numerical values 13.5 and 10 are variable: there would be two or three hundred different voltage and resistance values possible for this question. Third, weighted multiple-choice marking was introduced. Rather than one option always being marked right and the others wrong, different answers could be assigned different numbers of marks. Fourth, OASIS had always been able to offer students a large number of numerically different versions of the same question. However, previously these different versions always came with the same image. With further development of the software, OASIS was able to deliver different images as well as different numerical values within the same question.

During the first semester, the Department of Physics had carried out a small trial of OASIS with a group of their own year-one students. For this trial, two practice assignments of six questions each had been written, and the Department of Physics academics reported that they were pleasantly surprised at the popularity of OASIS. Following the success of this trial, they wished to use OASIS more extensively with their year-one classes. Unfortunately, they were not successful in securing funding to enable this to happen in 2003. The funding was declined with a statement to the effect that they should be using the University's software package, Cecil, not some other package. This statement sparked off some fairly acrimonious communications, in which the Head of the Physics Department and others pointed out to those in the university hierarchy that Cecil did not meet a number of practice and assessment needs that were satisfied by OASIS. For example, Cecil offered only multiple-choice questions, while OASIS provided numerically different variations for questions that were far more sophisticated. Further, even though Cecil was limited to one very simple question-type, it still had a slower response time than OASIS

The desire of the Department of Physics to use OASIS was seen as good validation of its potential. During 2003 I also discussed CAL and CBA in person with a number of overseas academics, and attended and presented a paper on OASIS at a top-rated conference (a list of relevant publications and conference presentations is presented in Appendix nine). These visits served to reassure me that we were on the right track. A number of packages clearly did not support learning and teaching as well as

OASIS. On the other hand, I did see a small number of software packages that offered better support than OASIS for some aspects of learning and teaching. There certainly were, and still are, ways in which OASIS could be improved; fortunately, the system has the flexibility to allow this. I saw three main areas where OASIS improvements would be welcome: reporting to academic staff on student activity and progress, discouraging student cheating during assignments, and uploading questions to the OASIS question database. It was decided that the third of these was the least pressing and could be postponed. The other two would be addressed in time for 2004.

Considerable interest and some excitement had been generated in the Department when I showed my colleagues the data collected by OASIS on student activity prior to an assignment. While most academics believe that students (being, after all, human) do procrastinate, there was little hard evidence to support such assertions. By contrast, the numerical data produced by OASIS on this matter were very clear. Unfortunately, it had not been easy to extract this data, and, in general, OASIS was still very reluctant to provide data. I considered that, relative to other software packages, OASIS was weak in this area. I was keen to obtain data on student activity and performance, and I knew that my colleagues would also appreciate ready access to such data. Unfortunately, providing this access would require considerable programming. While the programming was well within the capabilities of the academic charged with maintaining and developing the OASIS software package, he did not have sufficient time for such programming. Therefore I applied for and obtained a research grant that was, in part, used to hire a postgraduate student to undertake the programming. Part of this cost was borne by the Department. Most of the programming was carried out in the first semester of 2004. For this reason, the work is outlined in the following chapter, in Section 8.2.

In the area of discouraging student 'cheating' in assignments, I judged that OASIS and the way we used it were already relatively strong. For example, OASIS did use individualised questions. Therefore, if students wished to help each other, then they needed to explain methods rather than state answers. By contrast, some software packages relied on multiple-choice questions and allowed students several attempts at each assignment. However, it was clear from the year-one feedback that cheating

was an issue that bothered a significant number of students. There had been very little said on the matter by students during the year, or in the year-two, semester-one interviews and surveys: for example, there had been no mention of spreadsheets at all. The issue only became apparent at the end of the second semester when the interviews and surveys for year-one were carried out. Having received this feedback, I was keen to make OASIS more resistant to student 'cheating', and I considered that there were measures we could implement to achieve this. In preparation for their assignments, students generally worked through all the questions and quite often made notes on how to answer them. Sometimes students wrote out model solutions to each question. A few students even constructed spreadsheets to help them answer questions. Since the primary aim of the assignments was to promote learning, I was actually pleased with these sorts of activities. For a large number of the questions it would require considerable expertise to produce spreadsheets that could generate the answers. Anyone who could produce such spreadsheets certainly had an excellent grasp of the material. However, I was concerned to receive reports from some year-one students that some of their peers had been relying on spreadsheets produced by others. I judged that students who relied on the spreadsheets of others were unlikely to have spent much time mastering the practice questions themselves.

The fact that all the assignment questions were chosen from the practice questions raised another issue: the possibility that the students were not learning transferable skills but were simply learning how to do a particular set of questions. While I considered that test and examination results provided sufficient evidence of genuine learning, there were still some lingering doubts in the minds of a few of my colleagues. Furthermore, the fact that all the assignment questions were chosen from the practice questions may have encouraged students to focus on solving individual questions rather than thinking more broadly about their problem-solving repertoire.

Therefore, I decided to produce more questions for the two courses in which I was involved. These questions would be 'hidden questions', questions that were not available for student practice but could be used in assignments. In the courses next year, I planned to explain to students that some of the assignment questions would be questions that they would not have seen before. However, these questions would be similar to the questions they had already practised, so that, if they did understand

how to do the practice questions, then they should have no difficulty with the new questions. In this way I hoped to encourage the students to focus on their problem-solving skills, rather than to learn by rote how to do a set of questions. Also, I hoped that the advent of previously-unseen questions in the assignments would render the construction of question-solving spreadsheets less attractive. To further reduce the effectiveness of spreadsheets, I also decided to use the new ability of OASIS to provide alternative images (for example, circuit diagrams) within a single question.

Apart from the need to produce more questions for assessment purposes, there were other reasons why a substantial number of new questions needed to be written for OASIS. First, having taught and assessed my two courses twice, I now had a fair degree of insight into student strengths and weaknesses. There were certainly some areas where students had performed poorly in assessments and I wanted to strengthen these areas by targeting them with appropriate OASIS practice questions. Second, a number of existing questions in a range of courses needed to be rewritten in a minor way, either to correct mistakes in the answers, or to make clearer the wording in the question itself, or to make use of the new ability of OASIS to place variable values within text. Third, some instructors in courses that had not previously used OASIS were unhappy with the low skill levels shown by their students and wanted to lift these low levels by introducing OASIS practice questions into the courses.

Altogether, across my two courses and three other courses, there was a demand for 200 new OASIS questions. In order to meet this demand, the OASIS committee sought and obtained funding for $10 000. Two postgraduate students were hired to encode the questions. Of course, the success of this operation also depended on the ability and willingness of academic staff to write questions, provide adequate information to the students encoding the questions, and then check them once they had been uploaded to the OASIS question database. The reality was that not all staff did this to the extent indicated in their original requests. Since the funding had already been secured, the net result was that I made up the shortfall by producing more questions for my two courses than originally planned. Altogether, I wrote 60 new questions and produced 30 further questions that were based on modifications of existing questions. Once these new questions were added to the question database, the situation at the start of 2004 was as follows: my part of the year-one course

‘Electrical Engineering Systems’ had 54 questions for student practice and 39 hidden questions reserved for assessment purposes. In fact, several of these questions were hidden because they related to content I would no longer teach: I reduced the course content somewhat to allow more time for key concepts to be covered thoroughly. In my part of the year-two course, ‘Circuits & Systems’, there were 61 visible questions and 21 hidden questions. I considered that the question pool was now large enough to promote good learning in the areas to which OASIS was readily applicable. I also considered that using hidden questions in assignments would make them more valid, because more students would be encouraged to develop their own problem-solving skills, rather than to rely on others. It would be interesting to see what effect the use of previously-unseen questions would have on the assignment results.

I considered that the two action-research cycles of 2003 had been most successful. In terms of outcome validity, there was again a good level of evidence to suggest that student learning had been enhanced. This evidence came from a variety of sources (examination results, student surveys, and interviews with both staff and students) and was consistent with the previous year’s evidence. There still was no strong evidence supporting the other desired outcome, staff workload reduction, certainly not from my personal perspective. OASIS and its implementation were still in a state of change, and this was keeping my workload high. However, I did hope that, once reasonable stability had been reached, I would enjoy a reduced workload.

In terms of process validity, 2003 saw a significant increase in the rigour of the research. Valuable informal feedback from students was still received, but to this was now added the more formal feedback from dedicated surveys and formal interviews. Triangulation was further increased by involving others in the research: one colleague from CPD analysed the survey forms, while another from CPD interviewed staff members and wrote an independent report based on her findings. In this way, method and interpretation triangulation, as well as observation triangulation, were increased.

The use of interviews and surveys had provided students with better opportunities to influence the development and implementation of OASIS, thus increasing the action research’s democratic validity. The establishment of the OASIS Committee opened

up a clear channel for staff to have their say about OASIS. Staff interviews provided a further channel communication: all staff who ever had a significant involvement with OASIS were interviewed. Informal communications from both students and staff continued to have a significant influence too.

The action-research cycles in the coming year would naturally provide the opportunity to increase the validity of the research: for example, six sets of improved examination results would provide more convincing support for OASIS than four sets. It was also intended to obtain better access to the quantitative data collected and stored by OASIS, thus generating further possibilities for validating the research.

7.10 Summary

Prior to 2003, OASIS had run as a prototype software package on a shared server of low specification. In 2003, OASIS was implemented as a completely new software package on a more powerful, dedicated server. The new software was trialled successfully with one class in the first semester and then implemented fully in the second semester. A number of significant features, such as definable tolerances in marking, multiple images and variables in text, were also added to OASIS in this semester. The new software and server handled all loads admirably, including an assignment for 550 students in a single twelve-hour period. Usage figures indicated that OASIS was being used extensively for practice by most students.

Surveys and interviews were carried out and analysed for one class in each semester of 2003. These surveys and interviews indicated a good general level of student support for OASIS. Students perceived OASIS as easy to use and supportive of their learning. Test and examination results also indicated a high level of student learning. Further increasing triangulation, an academic from CPD carried out an independent investigation which confirmed the above and also suggested that OASIS was reducing staff workloads.

One key weakness of OASIS was the low level of statistical information provided to instructors. Considerable time and resources were put into an attempt to solve this

problem over the summer of 2003 - 2004. Unfortunately, this attempt was only partially successful, leaving more work still to be done in the future.

Interviews with year-one students had revealed that there was some use of spreadsheets during assignments. Partly to combat this, a large number of questions were written for OASIS over the summer of 2003 - 2004. The intention was that these questions would just be used in assignments. Thus, students in 2004 were to expect assignment questions that they had not previously seen.

Chapter 8. Mainstream OASIS: 2004

8.1 Purpose and outline of chapter

The previous chapter outlined the events of 2003, the first year the new version of OASIS was used. This chapter similarly describes the events of 2004; once more covering two semesters and therefore two cycles of the action research. Again, the first three research goals were addressed, the emphasis being on the third research goal and its focus on student learning. The year-two course 'Circuits & Systems' is the focus of the first cycle, described in Section 8.2. Changes from 2003 are outlined in Subsection 8.2.1. The most important change was the use of previously-unseen questions in assignments, this being the subject of Subsection 8.2.2. The student response to such questions was investigated through focus-group discussions and a survey, reported on in Subsections 8.2.3 and 8.2.4, respectively. The section concludes with reflections and plans for future action, presented in Subsection 8.2.5, foreshadowing the second semester and the second cycle for 2004. The year-one course 'Electrical Engineering Systems' provides the vehicle for this second cycle, described in Section 8.3, with changes to the previous year's practice summarised in Subsection 8.3.1. Unfortunately, the first assignment was disrupted by some serious University-wide system-performance problems, as described in Subsection 8.3.2. These problems led to an unplanned focus on the first research goal, as performance testing of the OASIS software was urgently carried out to identify and correct some software performance issues, as outlined in Subsection 8.3.3. These upgrades took place before the second assignment which, fortunately, ran smoothly, as noted in Subsection 8.3.4. Student interview responses and emails, presented in Subsection 8.3.5, confirmed and validated the use of OASIS as a learning tool but continued to question its validity for assessment. The findings of the 2004 research cycles are reflected on and subsequent plans for 2005 are outlined in Section 8.4. Finally, the chapter is summarised in Section 8.5.

8.2 Year-two, semester one, 2004

8.2.1 Changes from 2003

In order to more closely follow traditional marking practices, and to give students partial credit for correct working, consequential marking was implemented. Initially, such marking only dealt with different choices of units: for example, when a student specified an answer in millivolts instead of volts, the software would expect a numerical answer 1000 times larger. The software was subsequently extended fully beyond simple linear rescaling so that, if consequential marking was specified, a student who incorrectly answered one part of a question, but who then used her or his answer correctly in calculating a subsequent part of a question, would not be penalised repeatedly for the first mistake. A comment could also be displayed to the effect that, for example, the second answer, while not correct, was consistent with the first answer. Consequential marking did require extra calculations and increased the server load somewhat during marking; however this did not seem to be a problem.

As mentioned in Section 7.9, a student was hired to carry out some programming to facilitate easy access to the statistical data concerning student practice activities and assessment results. The work took place throughout the first semester. My role in this was to specify and confirm the programming outcomes, while another academic provided the student with the necessary technical support. It was hoped to add the following functionalities to OASIS:

(1) Statistics for all assessments. This would include the mean and the standard deviation for each question in the assessment and for the total.

(2) Statistics for individual practice questions. This would include the number of students who had attempted each question and information about the overall success rate for each question.

(3) A search facility for looking up individual students. On typing in the unique personal identity code for a student, the instructor would receive data concerning which questions the student had attempted, when the questions had been attempted, and the degree of success with each question.

(4) Statistics concerning the number of practice questions submitted by a class as a function of time. In this way, staff would be able to gain an insight into the student work-patterns on a day-by-day basis.

(5) Live-data displays during assessments so that staff could see which students had finished the assessment, which students were currently sitting the assessment, and which students had still to start the assessment. This live display would also show the results for students who had completed the assignment. Given the individual nature of the assignments, this would require a complete display of the questions given to each student, along with his or her answers.

This attempt to make the OASIS statistics more readily available was, for a variety of reasons, only partly successful. Items (1) and (5) above were fully realised but other items were, in general, only partly achieved. There were also issues with poor display and awkward navigational requirements. While considerable progress was made in this area, the reality was that still more was needed.

The above paragraphs detail general improvements to OASIS. Improvements that applied specifically to the year-two course "Circuits & Systems" were a result of my writing a large number of new questions for the course. These questions had been written partly to cover some areas of the course that had been poorly covered, and partly to produce a set of 'hidden questions' (previously discussed in Section 7.9) that would be used only for assessment. Some existing questions also had errors corrected or wording clarified. Once the 'hidden questions' had been separated from the practice questions, the next task was to order the questions logically. The final order for the practice questions was determined by topic order and by the principle that students should start with easier questions so they could develop the skills needed to tackle more demanding questions. The intention was that the use of 'hidden questions' in assignments would encourage students to develop their problem-solving skills in general rather than just focus on the practice questions. Some students produced model answers for all the practice questions, while others reported that colleagues had produced spreadsheets for the questions. I wanted students to realise that general problem-solving skills were far more important than being able to solve a particular set of questions. Further, I wanted minimise the number of students who relied on the efforts of others: those who, for example, would borrow a spreadsheet from another student. I also wanted to answer the few critics who still wondered if the students earned high marks only because the assessments comprised previously-seen questions. The fact that the students were

earning good marks in tests and examinations which comprised previously unseen questions seemed to be overlooked by these critics.

Apart from the introduction of some unseen questions in the assignments, no changes were proposed for the course. I considered that it had gone well and there was no need to change the content, the delivery, or the assessment schedule.

8.2.2 The OASIS assignments

As in 2003, there were four OASIS assignments and I was responsible for the first two, each worth 2% of the final grade. The first assignment ran on Wednesday of the second week of the semester while the second assignment was on Wednesday of the fourth week. Intentionally, for the reasons advanced in Subsection 7.4.2, these assessments came very early in the semester.

The students in 'Circuits & Systems' had met OASIS in 2003, but then their assignments had comprised only questions previously available for practice. I explained during lectures that the assignments this year would include some unseen questions and that, if the students understood how to do the practice questions, then they should have no difficulties with the unseen assignment questions. Prior to the first assignment, I also sent each student in the class an email which included the following sentence: "The questions in the assignment will be chosen from or similar to Q1 to Q16 (ELECTENG 202, Foundations DC) that you have already been practising". There appeared to be no significant additional anxiety concerning the assignment as a result of this announcement. As before, a number of students asked me for help with the practice questions, and I discussed a few of the practice questions during lectures. Prior to both assignments, the maximum student activity took place in the last 24 hours before the assignment, with fully one third of the activity occurring during that period. Before the first assignment, the average student made approximately 40 submissions to practice questions. In the period leading up to the second assignment, which covered more material, the corresponding figure was almost 60 questions.

During the first assignment I used for the first time the new live-data display: I could see which students had finished the assessment, which students were currently

sitting, and which students had still to start. The assignment ran from 8 am to 8 pm, and most students did the assignment relatively early in the day. Only four students out of 160 did not sit the assignment. During the assignment the software performed well with few problems: just two students experienced difficulties submitting their answers. Fortunately these answers were retrievable from the OASIS database, even though they did not appear in the live-data display. There were also a few similar problems with the second assignment, all minor.

Questions which had been available to the students during practice made up one third of the first assignment while the remaining two-thirds contained unseen questions. In spite of the inclusion of unseen questions, the results were as high as in previous years and the mean mark was 84%. The second assignment was based entirely on the 'hidden questions' and the mean mark was 83%.

8.2.3 Focus-group discussions

Previously I had interviewed individual students about OASIS. With this class I decided to facilitate, for the first time, focus-group discussions. I hoped that this new approach would reveal some fresh insights. At the least, I believed that bringing together a group of several students for a discussion might enable me to play a more minor role than would be possible in a one-on-one interview: I believed that the students might be less influenced by me and therefore more likely to speak frankly. I also considered that what was said by one student might spark off comment and discussion from others.

As described in Subsection 5.7.3, students were invited by email to take part in the discussions. Subsequently, two group discussions took place two days after the second OASIS assignment, just one day after the marks were finalised for the assignment. This tight scheduling was chosen because I wanted the discussions to take place as soon as possible after my last assignment and to be based on the way in which I, rather than another lecturer, was using OASIS. In the following quotes where more than one student speaks, the students are distinguished by S1, S2 and so on. The interviewer is denoted by CS.

The second focus-group discussion was longer and more spirited than the first, with considerable interaction between the students. At times I played an insignificant role

while the students debated issues amongst themselves. The first discussion group did not display much student-student interaction and in fact was similar in some respects to an interview. However, some useful information was volunteered by the students without prompting, including the following comment to the effect that the use of unseen questions in the assignments was an improvement:

> One thing I really think is better this year than last year is the testing. You get questions that aren't the same as the ones that you have practised directly. Not the exact same thing because it makes it more of test conditions, I think so.

The second group, while also voicing a preference for unseen questions in assignments, stressed that, if the motivational aspect of the assignments was to be preserved, it was important that these questions be similar to the practice questions:

> S1: Another thing I quite like is the relationship between the questions and the assignments. A motivation for me to do the OASIS questions is because they might be in the test... What I am saying is that it is quite good the way it has developed. If you suddenly make it that the questions are completely different from OASIS, then it just turns into a bit of a, just a test and all of a sudden, it is just like a test and you don't know what is going to be in it, you have no idea, and you have these written problems that you printed out and they just turn into problems like those that you have in your textbook. I don't know about the rest of you, but I am not very consistent with doing problems from the textbook. I find OASIS, for some reason, much more motivating to do, as opposed to the textbook.
>
> S2: All my assignments definitely motivated me to do the questions.
>
> S3: And if you know there are a couple of questions in there that might match the questions that sort of – I know it is probably not the best reason for doing it, but nonetheless it is still like a motivation to me. I did every single OASIS question....
>
> S1: I guess what I am trying to say is that you do keep some definite similarity between the assignments and the actual questions that you can practise, because if they diverge wildly – if I looked at them and said, oh they're textbook questions versus OASIS, I'd feel quite differently. Do you guys think it takes away a bit of motivation to do them or am I…
>
> S4: I think if there are similarities between the online practice and the test it's definitely a great deal of motivation to do online practice.

The motivational aspect of the assignments in general was frequently mentioned:

> S1: I think an assignment will make people work more than when there is nothing to push them.
>
> S2: So it's all priorities. If we've got an assignment for this we have to spend time on it...
>
> S3: It makes you practise pretty much every day. You want to finish all the questions in time, before the assignment.

Traditional assignments were seen as less motivating because the questions were known in advance and students tended to do just the assignment questions rather than a wide range. Traditional assignments also suffered from very delayed feedback:

> [In a traditional four-question assignment the students] are going to focus on those four questions whereas with OASIS they would think I have got to do a question on Thevenin [Thevenin equivalent circuits] , but no one question is going to be exactly what I need, so I would practise all the Thevenin questions. But, with an assignment, you would just do the exact question. Unless you cannot do it, then you have to find some easier ones. I think the worst thing about comparing a traditional assignment to OASIS is just the feedback that you get. By the time you get your results back from your [traditional] assignment, I find that I really don't care about why I got it wrong. But with OASIS, it is instant, so you know.

The students confirmed that they did use the feedback to sort out their mistakes. They also stated that, for traditional assignments in other courses, feedback usually was delayed by two to four weeks. Such delayed feedback was seen as of little use:

> S1: Unless it's the next day or two, or the next hour or two, by the time I've got it back, I've forgotten my working and I've actually got to look at the question and remember how I did it and take myself back to the original question – I've forgotten how I did it and it takes a lot of extra time to remember what my problems were and that sort of thing.
>
> S2: You know, we move on to another topic, so when we get the test paper back I can't remember…
>
> S3: It is better to know straight away, to know that you have got it wrong. So fresh in your mind, you go back to your working, as opposed to a week later... having to remember what the question was.
>
> S4: I think the main difference there is: it is better for you to find mistakes in your own working rather than wait a week for someone else to find mistakes in your working.

Practice on OASIS was seen as being valuable for improving problem-solving skills:

> S1: When I practise OASIS, now I look at the problem – I can know where to start and what to do and which is the easiest way.
>
> S2: Sometimes you are solving circuits one way and when you get to the end you realise you could have done it another way – a lot quicker. Which I've done quite a few times...
>
> S3: Because of OASIS, 202 is by far my strongest paper at the moment. I feel most up to date and most competent with all the work...
>
> S4: I found it very good last year, especially for getting the basics – because my basics were pretty weak after [high school]. It's some simple things like current divider and how to combine resistances and parallel circuits… It was really good to be able to brush up.

With respect to the above, the instant feedback and the ability to repeat questions with either the same or different numerical values were noted as important features:

S1: I like the way you get real-time feedback. You get to know the answers as soon as you have done the question, which is all I do, if I'm doing something if I'm by a computer, always do the question, and check the answer right away. Because I find that if I do a question and check the answer while I still have the working, it is really good to have real-time feedback.

S2: Another nice thing that I found is that I can go try again, if I want new numbers or I can go back and try with the same numbers, if I just want to do that last little bit. I am not sure that it is designed that way, but it is quite a nice thing.

The students considered that the key strength of OASIS was in motivating students and improving their learning. It was not seen as strong for summative assessment because it was possible for students to rely on others to improve their assignment results. However, in spite of the possibility of cheating, both groups appeared to be content with the current allocation of marks to OASIS assignments, and expressed a preference for unsupervised assignments rather than supervised OASIS tests:

S1: I think if you want to make it worth more than what it is, you would have to supervise. But at the moment, it is not worth much and if people want to cheat, they only cheat themselves because we have a lot more to do at the end of the year. But if we have done it properly, at the time, it's 10%. I think they are really valuable to you, just for your learning, and if they want to cheat for 2%, they can do it. I actually like the way you can sit it at any time during the day. I'd be pretty anti having to do it as a test.

S2: I think so too. If they really want to cheat that's ok but they will face the exam at the end of the semester. They will have to do it for themselves...

S3: I don't think it is a good idea to make it a fully blown test. I think if you do that you may as well just have a test written down on paper and it will be kind of taking it down to just a standard university test and I think OASIS is a lot better than that.

S4: I mean the assignments are open book anyway, so it's not really a test of your ability. It's not the same conditions as an exam.

S3: It keeps it conceptual. I don't have to remember what complex capacitance is... I just have to understand how it all works and the general ideas and concepts behind it – maybe I can't remember the formula, I can look it up, and that sort of thing. It makes it a lot better. That makes you concentrate on the idea behind the questions, as opposed to specific little nit-picky formulae and stuff like that.

S5: Yes. That's one of the strengths of it as well.

It was suggested that, in order to neutralise the advantage gained by students who cheated with their assignments, that the assignment mark could be limited so that it could not improve a student's overall mark by more than a certain percentage. Thus, a student who scored 100% on the assignments but only 20% on the final examination would not be credited with the score of 100% from the assignments but would be credited with an assignment score closer to 20%:

S1: could you do something like the OASIS grade cannot increase your exam grade by a certain percent - internal cannot bump you up by more than 10% of your exam...

S2: I quite like that. There are a lot of people who compare answers for assignments… But if you say, yes right, this is what happens... if you cheat on that, it's not going to benefit you. Do the questions and know the work and then you will do well in both [assignments and examinations]...

S3: Ultimately, at the end of the day your exam mark will decide the grade for the course. If everyone does well in OASIS, that's to be expected. People who have cheated, they get 90% for OASIS, but they still do badly in the exam and I don't think it's going to affect final standings.

S4: And if you find it does, you can still do that restriction of the difference or something.

Some students expressed a wish to have hints provided with the questions, but such wishes tended to be expressed with some reservations:

S1: Yes, I think it would be good if there could be several hints for each piece, if we couldn't get it, but then it is good to talk to other people about the questions as well.

S2: I suppose once you get used to getting hints… it would be better if you had skills or people that you worked with and were able to talk to them about it, because that would be a lot better than just popping up the hint.

In the second focus group there was a particularly spirited discussion on the issue of whether hints should be provided or not:

S1: Some of the things that I do, I keep getting wrong: it would be nice to just click on a hint button. It could be abused, as most things, but…

S2: I know, personally, I wouldn't use it until I have tried and got the answer. If I get it wrong, try again and if it is still wrong then… maybe give me a hint with the answer.

S3: The reason why I don't like the idea is that because one of the cool things about OASIS is that I have learnt is to look at the problem and say – right, I go through it in my head – all the different ways of doing it and looking at it and say, right – I can do it that way. Like if you have a circuit as an engineer you are not going to have a hint that says – use

mesh-current... that's like three-quarters of the game done. Just plug in the numbers and that's done. That's a bit silly.

S2: That's a good point. But if the hint was only given with an answer, possibly only if the answer was incorrect…

S4: I missed the first couple of lectures and I missed the lectures on delta-wye transformations and looked in the book and I saw what it was, but I missed the example, when you actually used it. I looked at the OASIS questions on it, I think 14, 15, and 16 are on it, and actually I was completely clueless. I asked people the next day "is this delta-wye?" and they said "yes" and it was much easier. So it is really something like that – when you put the incorrect answer or something.

CS: The way it operates at the moment, which is a bit slow, is that if you get stuck on a question, you e-mail me, then I write an answer or say "try this" or something like that.

S5: Yes, but as we said, it takes a lot longer and gives a lot more work for you – responding to 15-20 emails. It might be easier if you put it as part of OASIS…

S2 If you only give the hints with the answers and someone does it and knows how to do it, they can ignore the hint or they won't even see the hint – that does not really matter. It is only for people who have tried it and got it wrong.

S1: I don't see why you have to be bothered with the fact that people might cheat: it shouldn't concern you if they to do that, that's up to them.

S2: That's my kind of thinking – if people want to abuse it, that's their problem…

S5: Hints are just a time-saving thing. Because if I don't get it right, I've got to note it down and ask someone at Uni or ask you or something like that. If we could look at the hint, it would be so much better.

S2: The other thing is when you get home, you have some circuits that you don't know the answer to, and you don't have hints and you don't have people to ask, because no one has ever done it. You could even have made up the circuit yourself. I don't know how that would work but…

S5: I think in that case you'd have to be able to ask a colleague…

S2: Yes, which is like practising by doing as a year-two student. I can imagine that, as a fourth-year engineer, you'll be solving circuits that someone gives you or you make them up yourself.

Several points are made in the above discussion. The students commented that, although they themselves would only access a hint after they had made considerable effort, they felt that hints could be abused by some other students. It was later suggested that such students should be of no concern to those correctly using the hint system. However, such students should be of concern to the instructor because, by accessing the hints, the students could be bypassing some very important steps

needed in the development of problem-solving skills; for example, the ability to consider a range of problem-solving techniques and select the most appropriate. This point is made by student S3 who suggested that being told what problem-solving technique to use is "like three-quarters of the game done". It was noted that hints could be particularly useful for students who have missed part of the course. It was also noted that hints speed up the process for students who have reached an impasse. Against hints, the argument was made that more advanced students and practising engineers should not always expect hints because some problems are new ones that have not been considered previously. From this it follows that, beyond a certain level, the provision of hints in the education of an engineer becomes unrealistic and counterproductive.

It was also noted that it was difficult to find questions of a particular type because the questions were simply numbered and descriptive lacked labels. This had been done because such labels would prompt certain solution methods. They would therefore detract from the development of problem-solving skills and could also be misleading because there is more than one technique for solving most problems. It was clear from the discussion that the students appreciated these points:

S1 One criticism of OASIS – a fairly minor one – a lot of the time, someone who… just wants to go back and repeat a certain section of the problems, you go to a list of 49 problems or whatever and it is difficult to remember which ones deal with superposition… So it might be a good idea to categorise them…

S2 If you do it that way – if you say for one question, you could say superposition or whatever, then the people who are doing it might already get how to do it from the name.

S3 You could split them up into something like general and sub-topics like attenuators, say, that can easily be categorised. But when you have something like circuit-solving techniques, then you can do a problem with supernode or mesh or superposition. It doesn't really matter which way you do it.

S1 Practically there might be a few issues there, it's just a thought…

S4 It is done in fairly logical order, which is nice. It progresses and there are a couple of questions on this and a couple of questions on that, which is good and helps you to find it.

Overall, the students were supportive of OASIS and its implementation. Given that all the students had volunteered for the discussion groups, this was not altogether surprising. However, I did derive many useful insights. In particular, these students

were the first to experience unseen questions in OASIS assignments; it was pleasing that they approved of the use of such questions in assignments. Further, while student surveys had been heavily in favour of hints, the focus-group discussions were far more ambivalent on that topic and generated some powerful arguments against them.

8.2.4 The OASIS survey

The survey conducted with the year-two class 'Circuits & Systems" in the first semester of 2004 was the third time the same OASIS survey had been conducted. However, the previous two surveys had been conducted in lectures and there was the possibility that the survey results were not truly representative. For example, perhaps only those students who had a strong opinion about OASIS were sufficiently motivated to complete and return the survey. For the third survey, a different approach was used with the intention of obtaining a clearly representative sample.

The forms for the third survey were given to students at the start of the final laboratory-interview session, conducted near the end of the semester. I explained to the students that I was keen to receive their feedback about OASIS and that I would be grateful if they could complete the survey form while waiting for their interview and give the interviewer their completed anonymous survey form at the start of the interview. This way, the students had ample time to complete the survey form. Further, since the laboratory interview was not regarded as a particularly threatening experience, it was considered that the students would give adequate attention to completing the survey. Not all students were asked to complete a survey form but, from the large and representative group that was asked, there was a return rate of virtually 100%. In total, from a class of 166, 125 students completed the survey.

As outlined in Subsection 7.4.3, the survey form consisted of two pages. The survey forms were, as usual, analysed by CPD and then returned to me so that I could read the comments on the free-response page. The CPD analysis of the first page of the survey is shown in Appendix eight. The data from this first page is summarised in Table 8.2.4. The timing of the survey was not ideal: I would have preferred it to have taken place at the conclusion of my set of lectures, rather than some weeks later, because the lecturers who followed me in the course did not use OASIS as much as I did. In spite of this, the survey results were definitely encouraging. Most students

(89%) agreed or strongly agreed that OASIS was easy to use, while only 2% did not. These figures were very similar to those returned by the survey of the year-two class in 2003 (91% and 1%, respectively). The instant feedback was appreciated, with 82% agreeing or strongly agreeing with the statement "I like the instant performance feedback using OASIS". The corresponding 2003 figure had been fractionally higher at 86%. 80% agreed or strongly agreed with the statement "OASIS helped improve my skill level". This result was slightly down on the 2003 figure of 84%. Continuing the trend, 76% agreed or strongly agreed with the statement "The problems provided in OASIS helped me understand the course material better". Again, the 2003 figure had been a little higher at 80%.

Table 8.2.4

Student evaluation results for OASIS in year-two classes, semester 1, 2004

Item	**SD** (%)	**D** (%)	**N** (%)	**A** (%)	**SA** (%)
OASIS is easy to use	0	2	6	46	43
I often use OASIS	2	13	35	35	14
OASIS helped improve my skill level	0	5	14	46	34
I am more confident about my learning after using OASIS	0	4	13	54	27
The problems provided in OASIS helped me understand the course material better	0	2	18	50	26
OASIS helps me to prepare for the assessments	1	4	11	50	33
I come to campus less often because of OASIS	15	34	30	8	2
I like the instant performance feedback using OASIS	1	1	14	40	42
I prefer to do problems from the text rather than OASIS problems	6	25	46	14	6
It would be a good idea to have OASIS in other courses	2	2	13	38	40

(SD: strongly disagree, D: disagree, N: neutral, A: agree; SA: strongly agree)

However, against the trend, 81% of the 2004 students agreed or strongly agreed with the statement "I am more confident about my learning after using OASIS" while the

corresponding 2003 figure had been only 74%. In general, the 2004 year-two survey results, while still showing a most positive student perception of OASIS, were slightly less positive than the 2003 results. The clearest difference in the two survey results was in response to the statement: "I often use OASIS". For the 2003 group, 67% of the students had agreed or strongly agreed with this statement. For the 2004 group, the corresponding figure was 49%.

However, although the 2004 survey results appeared, overall, to be slightly less positive than the 2003 results, I did not believe that the 2004 class as a whole was less positive about OASIS than the 2003 class. Rather, I believed that the above differences were due to the differences between the two groups that completed the survey forms. In 2003 there had been a 40% completion rate for the survey forms, handed out during the last lecture of the semester. In 2004, by contrast, there had been virtually a 100% completion rate for the survey forms. The statement "I often use OASIS" produced the largest discrepancy, with the surveys indicating much less OASIS use in 2004. However, the reality was different, and the data recorded by OASIS itself indicated that OASIS usage in the two years was very similar. Given this fact, the difference between the responses to the statement, "I often use OASIS" suggests that the students who completed the survey in 2003 tended to be those who used it more than average. Such students would be, most likely, more enthusiastic about OASIS than the average student. This difference would certainly adequately account for the slightly less positive survey result in 2004.

Actually, the similarities between the 2003 and 2004 year-two surveys outweighed the differences. Even though the return rates were 40% and 100%, respectively, the results for almost all items were only marginally different. Perhaps it may be concluded that, when a group of students is asked to complete the OASIS survey during a lecture, the group that actually does is approximately representative in most respects (although not in terms of how often they use OASIS).

In addition to the 10 items set out in Table 8.2.4, the survey also contained five free-response questions. When asked, "What do you like most about OASIS?" by far the most frequently mentioned feature was the instant feedback. This had also been the most common response, and by a large margin, to the question in the previous two

surveys. One student saw it slightly differently, writing the following response to this question: “Gives you bad marks instantly”. Students also frequently mentioned that there was a large variety of questions and questions could be repeated many times with different numerical values each time:

> It marks it instantly and there are lots of questions to practise.
>
> Variety of examples, types of problems...
>
> The levels of difficulty of questions and the variety of questions...
>
> Being able to try problems with different values allows better understanding.
>
> I can practise questions with different values every time so I really know how to get the right answer when the values are different.

In response to this question, some students commented on how OASIS helped them improve their skills, perhaps anticipating question four:

> The chance to sharpen up my basic skills...
>
> It is a good utility in improving problem-solving skills...
>
> Instant answers. You can work backward if you don’t fully understand. This helps in the understanding of solving techniques.

When asked “What do you like least about OASIS?” students most frequently noted that answers are provided without explanation. This had also been the most common response in the previous two surveys. Student responses to this question included:

> Do not show the method of obtaining the solutions.
>
> No explanation or hints for difficult problems.
>
> If we have no idea how to do a problem, there are no resources or hints to help you.

One student commented here that:

> It can get frustrating if the question is really hard! But the lecturers are helpful when I need help with the questions so that is good.

Although students had praised the variety and number of OASIS questions in their responses to the first question, here in response to the second question they sometimes commented on the lack of question variety and coverage. The topics mentioned as needing more coverage were not in my part of the course, sometimes not in the course at all.

The most common response to the survey question, "What changes or improvements would you suggest for OASIS?" was a request for more questions and/or for more topics to be covered. One variation on this theme was a request for:

> An instant circuit designer. So you get different circuit designs, not just different numbers.

The second most common response was a request for worked solutions to the questions, a request that had been the most frequent response to this question in the previous two surveys:

> Worked answers for the harder questions perhaps.
>
> Explain some of the stuff, instead of us spending ages figuring it out.
>
> A link to refer to our notes.

A few students stated that they wanted more questions that addressed concepts rather than simply required the application of a formula:

> More questions. Questions which are concept-based rather than straight formulae application.
>
> More questions; questions which test your understanding rather than just applying formulas to get the answers.
>
> Instead of just numerical, maybe it could be revolutionized and start asking conceptual (wordy) questions to test understanding.

Pleasingly, a few students stated that no changes or improvements were needed: "nothing", "can't get better".

By a considerable margin, and following previous responses, the most common replies to the question "What benefits do you think students get from using OASIS?" centred on the idea that it provides students with good opportunities to practise problem-solving and improve skills:

> A lot of practice and new knowledge.
>
> Improvement to their skills, more understanding.
>
> Students can keep practising the same question until they get the hang of the question. And they can practise with different values every time.
>
> Quite a lot [of benefits] actually. One is from doing a lot of problems. Thinking about the problems for themselves.

Related to this was the idea that students could get feedback about their progress and level of skill and understanding:

> Instant feedback on where they are (what level they are at).

> They are able to see what concepts they find difficult.
>
> They can test their understanding of what is being taught in the lectures.

A few students mentioned the motivational aspects of the OASIS assignments, and that the assignments encouraged them to keep more up to date with the subject:

> Forces them to revise topic, giving easier time studying for exam/tests.
>
> Their grades would increase.
>
> The assignments mean that I am more likely to practise the problems and way before exams too!

As had been the case for the previous surveys, there was little pattern to the responses to the final question: "Have you any other comments to offer about OASIS?" As before, the most common response to the question was an encouraging and often fairly general statement to the effect that OASIS was useful and/or good:

> It's a good idea, helps us a lot in many areas of study.
>
> Sell it, it's a good system. Maybe include a 'how to solve' section.
>
> Very good way to learn using modern technology.

Some themes that had been mentioned previously by the students surfaced again:

> Have small and frequent assessments, prevent people from slacking off during the semester.
>
> I think my strong areas in electrical are those covered by OASIS in assessments. Those not covered are the ones I'm struggling with.
>
> OASIS greatly helped me in my understanding.

Interestingly, this last comment was written by a student who had, in response to an earlier question, written "there are not enough questions to test my understanding on"!

Overall, the feedback from the survey was positive. As in 2003, students reported that they used it frequently, found it easy to use, and felt that it helped their learning. Suggested improvements included coverage of more topics and a greater variety and number of questions. Again, a number of students stated that they wanted OASIS to provide model answers, or at least worked solutions, to the questions. I was still reluctant to provide these, for the reasons outlined in Subsection 7.4.3.

8.2.5 Reflections and plans for future action

There had been few changes from 2003 in terms of the software and its implementation in the year-two 'Circuits & Systems' class. OASIS, I considered, had been well received that year and the learning outcomes had been good. I was confident that it would be the same this semester, and the evidence did confirm this. Again OASIS was frequently used by most students, and both survey and interview results indicated that students found it easy to use, appreciated it as a learning tool, and believed that it helped them learn the material and improve their skills and understanding. I had received many emails from students about OASIS questions and this had helped me modify some of my lectures to address student concerns.

One key change this semester had been the shift to assignments that included previously unseen questions. The aim here was to ensure that students really had mastered the material: it had been suggested that OASIS was simply teaching students to solve a set of problems rather than helping them develop more general problem-solving skills. Two-thirds of the first assignment consisted of previously unseen questions. The mean mark was 84%. The second assignment consisted entirely of 'hidden questions' and returned a mean mark of 83%. These assignment marks were every bit as good as those gained the previous year when the assignments had consisted entirely of questions that had been available for student practice, providing strong evidence that OASIS was helping the students develop problem-solving skills, not just teaching them how to solve a limited set of questions.

In the final examination the two questions on my part of the course (which was well supported by OASIS questions) returned averages of 84% and 77%. These averages were higher than the averages for the other four questions which ranged from 31% to 63%. Colleagues had reviewed my examination questions prior to the examination and deemed them to be, if anything, too demanding.

The above results were most pleasing. The assignment results maintained the previous year's high level though the great majority of the assignment questions were previously unseen questions. The examination results for the two questions based on my part of the course were a useful improvement on the previous year's results (2003 average 73%, 2004 average 80%). Together, these assignment and

examination results suggested that the students had performed better in 2004 than in 2003. This was a most gratifying result, particularly given that, in order to maintain class numbers, the course entry requirements had been slightly relaxed and, therefore, a decrease in performance might have been expected. Such a decrease was indicated by the other four questions in the final examination: in 2003, the average for these four questions had been 61% while in 2004 the average was 51%.

This semester I had ensured that a large representative sample of the students completed the OASIS survey. While the results of this survey were interesting to peruse and supportive of OASIS, I considered that they did not reveal any significant new information. However, the fact that the survey produced results similar to previous surveys suggested that an adequate level of feedback could be obtained simply by distributing survey forms during a lecture and collecting in those completed. The only clear area of discrepancy appeared to be the students' perceptions of the extent of their OASIS usage. However, in this area the OASIS software itself was able to collect data directly. Given the considerable time involved in administering and analysing the surveys, and given that the three surveys conducted to date had produced consistent feedback from students, I judged that there was little to be gained by conducting further surveys and, instead, the time could be more productively devoted to other aspects of the research programme. Surveys could be conducted again in the future if events, such as significant changes to OASIS or the way in which it was implemented, were deemed to warrant it.

By contrast, I considered that the focus-group discussions had yielded a wealth of pertinent information. In particular, the students in the second group had initiated a lively debate on the provision of hints. This matter had been explored in some depth, with the discussion debating how extensive the hints should be, how readily available they should be, and whether hints might detract from the problem-solving experience. It was also noted that the provision of hints is somewhat unrealistic because hints are not available in the real world where engineers typically work on new problems rather than problems that have already been solved.

The new version of OASIS had been implemented across all participating courses half-way through 2003, and I was now in a position to reflect on how successful it

had been over a twelve-month period. As outlined above, I certainly did consider that the implementation had been successful. However, its success was producing a few potential problem areas. It was now becoming more of a mainstream teaching and learning aid, and a larger number of courses were not only using it but also starting to rely on it. Whereas previously it might have been offered in a large course of 500 or 600 students only as a practice tool, it was now shouldering a significant assessment load. Further, whereas previously it had been used only within our own Department, now the Department of Physics was also using it with some of their own very large classes. It was becoming clear that, in order for OASIS to properly play its new role, two issues needed to be addressed. First was the issue of increasing student numbers using OASIS. In the past we had considered that being able to handle around 500 concurrent users was acceptable; now we were confronted with the possibility of twice that number. For example, one large class could be practising on OASIS while another was completing an assignment. It was important to check the ability of OASIS to handle these larger loads before they became a reality. Furthermore, it made sense to check the software itself for any areas that created bottlenecks and limited performance. Consequently, two year-four students were given the task of running performance tests on OASIS and attempting to identify any performance bottlenecks. This task was done as a research project, a required part of the Electrical Engineering degree programme. It was supervised by one other academic and myself, and was to be completed by the end of the year.

The second issue that needed to be addressed was the long-term viability of OASIS. The OASIS software itself had evolved rapidly and was still undergoing regular change. Knowledge of the software and its implementation resided with a very small number of people, and was largely undocumented: the departure of just a few people from the Department could have rendered the software useless. As OASIS was becoming more widely used, and as courses were making longer-term commitments to it, adequate documentation was becoming more critical. The task of producing suitable documentation was taken on by the OASIS committee. Two different types of documentation were deemed necessary. The software itself needed documenting, and this task was shared between a committee-member and a postgraduate student. A user's manual was also deemed necessary. Another postgraduate student, one who had encoded many questions and who was very familiar with OASIS, was employed

to carry out this task, overseen by another committee-member. The intention was that these items of documentation be completed and available by the end of the year.

The commercial future of the software was another issue that needed to be considered in the formulation of any longer-term plan for OASIS. There was some discussion about whether OASIS should be marketed, perhaps as part of a larger package, whether it should be made open-source, or whether the status quo should persist. No decision was made at this time although some interest had been shown by UniServices, the commercial arm of the University, in marketing OASIS.

8.3 Year-one, semester two, 2004

8.3.1 Changes from 2003

There had been major changes in the year-one course 'Electrical Engineering Systems' and its assessment in going from 2002 to 2003. By contrast, there were very few changes in going from 2003 to 2004. Rather, the period was one of consolidation. The 2003 course is described in Subsection 7.6.1. The content of my part of the 2004 course was virtually unchanged; just a fractional trimming of the content and a few more examples and demonstrations. The assessment was identical to that of 2003; however, I did intend to use some previously unseen questions in my OASIS assignments. Although previously-unseen assignment questions had already been implemented in my year-two class with little noticeable effect, it was the situation reported in the year-one class that had motivated their implementation (see Section 7.7). Therefore there was considerable interest in discovering what effect this change might have. The use of previously-unseen assignment questions was the only significant change for this cycle of the action research; the emphasis would be on collecting data to confirm earlier findings rather than implementing change. The student body taking the course in 2004 was also the same in size and academic achievement (based on high school results) as the 2003 student body.

8.3.2 The first OASIS assignment

As in 2003, just two OASIS assignments took place: since only my part of the course was well covered with OASIS questions, only my part of the course could be assessed with OASIS assignments. Unfortunately, my lectures did not start till week

four of the semester, so I was unable to use assignments as a way to provide early performance feedback. The material of the first three weeks was assessed by a conventional test in week four, but the test results were not returned to students until the second half of the semester. Consequently students did not receive a significant level of objective information about their achievement level in the first quarter of the course until more than half the course had passed.

The first OASIS assignment was held on the Friday of week six of the semester, the last day before the mid-semester break. Unfortunately, the assignment did not run smoothly. Major problems appeared around 1 pm when students started complaining that the system had slowed down dramatically. The problem was not restricted to OASIS: the entire University network was running slowly and the University's learning management system, Cecil, stopped functioning altogether. OASIS was still functioning, but only just. A number of students could not complete the assignment in the allocated hour. To get round this problem the assignment was extended by 24 hours till 8 pm Saturday. This arrangement did not suit all students; some had planned holiday travel. Such students completed a paper version of the assignment. Since the University network was having major problems, and Cecil was inoperable, it was difficult to contact students electronically about the new arrangements. With a class of 560 students, the result was that I spent two or three hours explaining matters to students who visited my office to complain about the network and/or the OASIS assignment. Fortunately, the students were most supportive once I reassured them that their situation would be handled as fairly as possible. The response-time difficulties with this assignment motivated a number of performance upgrades which are detailed in Subsection 8.3.3.

There was also a problem with one assignment question. For this question the answers stored in the OASIS question database had been rounded to four decimal places, as the answers had been for many other questions. However, for this particular question, the answers were sometimes rather small and that led to problems when marking to the default 1% tolerance. For example, suppose a particular version of the question had an exact answer of 0.00214. This would be rounded and stored as 0.0021. However, if a student then submitted the correct answer of 0.00214 this would be compared with the stored answer of 0.0021, judged

to be outside the 1% tolerance, and marked wrong. In order to correct this problem, I spent most of one morning checking each student's answer for the troublesome question. In the class of 560 there were about 50 students who earned one or two more marks for the question. This problem was not inherent in OASIS, and could be solved simply by rounding stored answers to four significant figures rather than to four decimal places.

I had introduced the students to OASIS soon after I started my lectures, in week four of the semester, when there were also a number of other demands on their time. Perhaps this partly explained why there was only a relatively small amount of OASIS activity until the last day before the assignment. Further, the placement of the assessment on the last day before the mid-semester break was certainly not ideal either: I was asking the students to prepare for yet another assessment at a time when they were likely to be mentally exhausted and looking forward to a two-week break from their studies. The previous year I had located the OASIS assessment in week seven of the semester, just after the mid-semester break. That timing had had the advantage that the students would be likely to approach the assessment with more mental energy, but also the disadvantage that the students would not receive any feedback from assessments till nine weeks after the start of the course. This year the students had received feedback three weeks earlier, but the cost was an assessment on the last day before the two-week break.

Prior to this assignment, the students had submitted a total of 22000 answers to OASIS questions (approximately 40 per student), with approximately two-thirds of these submissions taking place in the last 24 hours before the assessment. This was significantly fewer than the previous year's figure of 27000 submissions (approximately 50 per student). However, most of this difference was, I judged, due to the fact that the previous year some students had made good use of the time afforded by the mid-semester break to work their way through the OASIS questions. The number of answers submitted in the last 24 hours prior to the assessment was around 25 per student for both 2003 and 2004. These figures did not actually capture all the student activity on OASIS: it was known from interviews and informal discussions that a number of students printed out the questions and then worked through them on paper two or three times. This activity could not be recorded by

OASIS. Balancing this was the likelihood that some students submitted answers with little thought or calculation: to obtain the correct answers, for example. However, while the numbers could not be perfect indicators of the amount of mindful, problem-solving activity, they did reveal activity patterns. It seemed that, in a busy week, most students had managed assessments serially, starting serious work for each only after the previous one had been completed. Further, most students had not planned ahead for the busy periods by starting preparatory work earlier when there were fewer demands on their time.

The OASIS data above make it clear that the class of 2004 had, owing to the unfavourable timing of their assignment, prepared less well for it than the class of 2003. As expected, this poorer preparation resulted in poorer assignment results. In 2004 the average mark was 75%, with 125 students gaining full marks and a further 89 students gaining 12 out of 13. In 2003 the average mark had been 82%, with 200 students gaining full marks and with a further 87 students scoring 15 out of 16. However, a clear comparison between the two years is difficult. First, it is likely that the network problems associated with the 2004 assignment adversely affected a large number of students: they would have been worrying about whether the system would allow them to complete the assignment in time, not focusing fully on answering the questions. Second, the 2003 assignment consisted of previously-seen questions whereas only one of the six questions in the 2004 assignment had been seen before, and I had included that question precisely because it had given the students particular trouble during practice (in fact the class scored more poorly on that question than on any of the previously-unseen questions). While the change to previously-unseen questions had not affected the performance of the year-two class, it may have affected the performance of the year-one class: it is likely that the last-minute preparation carried out prior to the assignment would have focussed on being able to do the practice questions rather than on developing general problem-solving skills.

8.3.3 OASIS performance upgrades

While the lack of system speed during the first OASIS assignment had been a University-wide issue, it was important to ensure that OASIS itself was not contributing to the problem. To this end, soon after the first assignment, the OASIS server's RAM was tripled to 1.5 GB. The software too came under considerable

scrutiny. The two year-four students who had been carrying out performance testing established that the software was executing some tasks relatively slowly. Further, in the middle of the year, the OASIS programmer stated that he considered the software to be running as slowly as ever. This slowness was likely to be the result of many ad hoc changes to the software, often carried out in some haste to resolve problems or meet staff requests. In the second half of 2004, the OASIS programmer worked in conjunction with the two year-four students to identify and eliminate performance bottlenecks (Beh, 2004; Ng, 2004). The two students conducted trials of OASIS on a test server, a machine much slower than the production server. When these two servers were given the same simulated task for a range of user numbers, their response times became quite different with increasing user numbers, as shown in Figure 8.3.1. A dashed horizontal line marks a response time of three seconds, the maximum response time deemed to be acceptable.

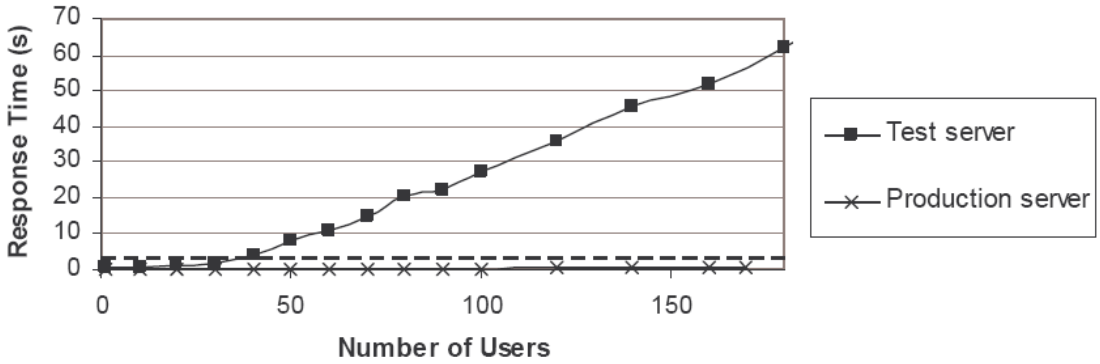


Figure 8.3.1: A comparison of test-server and production-server response times as a function of number of users. Adapted from Figure 6.8, Beh (2004).

The performance trials were conducted on a test server rather than the actual OASIS production server because the latter was continually used by students for OASIS practice and assessment. Therefore it could not be used in trials that were designed to drive it to its performance limit; at such times students would have found its performance unacceptably slow. The test server, being lower-specification, also had the advantage of revealing performance shortcomings in the software more readily.

The investigations revealed that the loading of questions was particularly slow. There were two main reasons for this. First, slow image-retrieval, particularly for large

images. Second, accessing a question involved retrieving more data than was strictly needed: all numerical variations of the questions were retrieved rather than just one.

Partly to address the first issue above, and partly to avoid the payment of licensing fees, al images drawn in the Graphics Interchange Format (GIF) were changed to the Portable Network Graphics (PNG) format. In some cases this led to a three-fold or better reduction in file size. Trials on the test server showed that even a modest reduction in file size would lead to considerable performance gains. The effect of question file size on server response times is illustrated in Figure 8.3.2. Again, a horizontal line shows the longest response time deemed acceptable, namely three seconds.

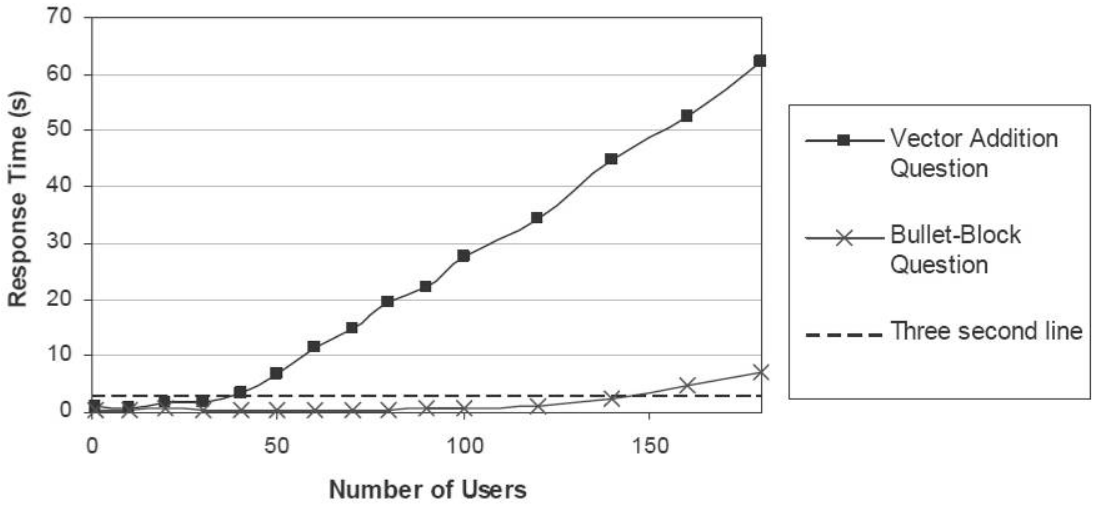


Figure 8.3.2: The effect of file size on response time for the OASIS test server. Adapted from Figure 6.2, Beh (2004).

Two different simulated assessments were trialled. The first was based on a vector-addition question that contained five images. For this assignment, the server had to transfer a total of 189000 bytes during each simulated user session. The second simulated assessment was based on a conservation of momentum and energy question (the 'bullet-block' question) that contained two images. For this assignment, the server had to transfer a total of 148000 bytes during each simulated user session. Physics questions were chosen for these trials rather than electrical-engineering questions because the former tended to have larger graphics associated with them.

It can be seen that the test server could only deliver the assessment with the larger graphics to 40 or fewer users while maintaining an acceptable response time, while for the assessment with the smaller graphics the corresponding number of users was 150. For this number of users, the former assessment had a response time of 50 seconds, the latter, 3 seconds. It is clear that even a modest reduction in image file size and, more generally, the amount of data transferred, had a profound impact on response time.

In order to further reduce the amount of data required to be transferred from the question database, a way was found to effectively cache static images. With this upgrade, static images, such as the OASIS logo, did not need to be retrieved from the question database each time a new question was displayed. The effect of this upgrade on the ability of the OASIS test server to deliver a simulated assessment based on the vector-addition question is shown in Figure 8.3.3.

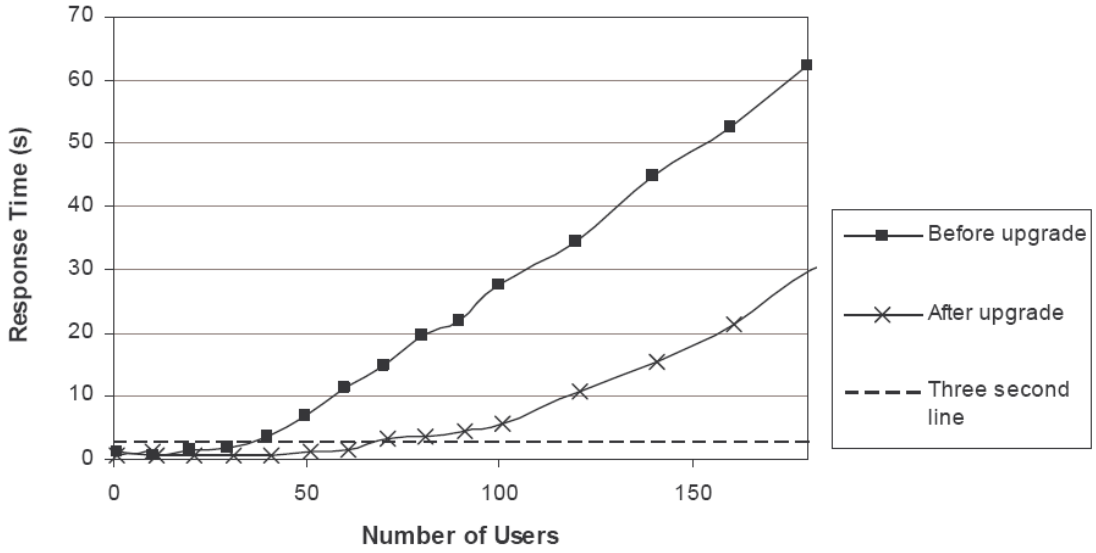


Figure 8.3.3: The effect of the static-image caching upgrade on response time for the OASIS test server. Adapted from Figure 7.2, Beh (2004).

It can be seen from this graph that the caching of static images allowed the number of users to increase from 40 to 70 before the response time became unacceptably slow. For 70 users, the test server's response time was 15 seconds before the upgrade and just three seconds after the upgrade. This upgrade also decreased data-storage requirements: each static image needed to be stored just once in the cache.

A major software upgrade that addressed several performance issues took place in September 2004. The main issue addressed was the fact that each time a student attempted a question all its numerical variations were retrieved from the database. After the upgrade just one numerical variation was retrieved when the question was presented. The effect of this upgrade on the performance of the OASIS test server for the simulated assessment based on the vector-addition question is shown in Figure 8.3.4. The improvement was radical. With the upgrade the response time was only 0.6 seconds, even for 320 simulated users, the maximum number in the trial. Prior to the upgrade, just 40 concurrent users had produced the maximum permissible response time of three seconds.

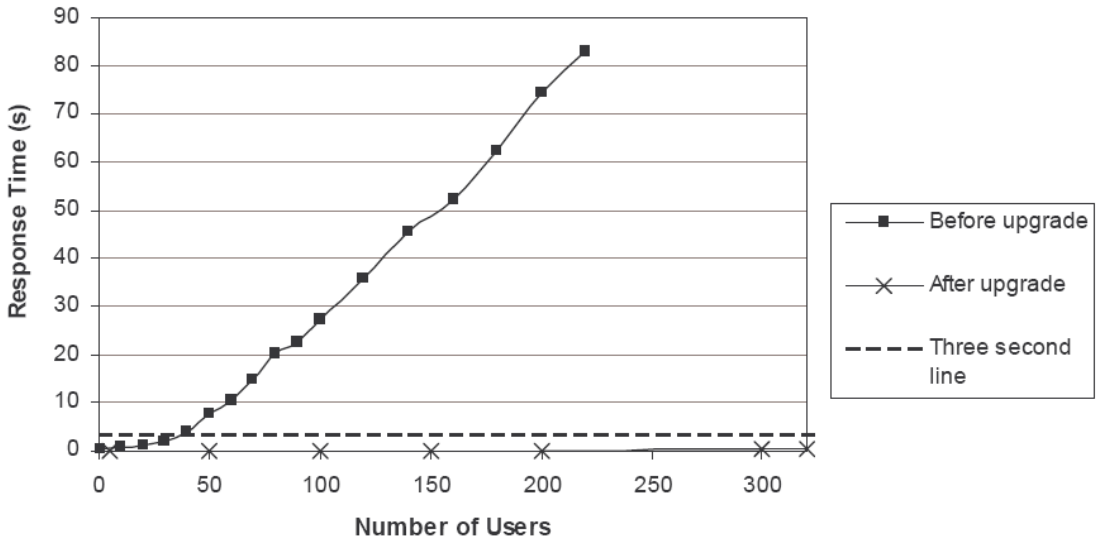


Figure 8.3.4: The effect of the question-retrieval upgrade on response time for the OASIS test server. Adapted from Figure 7.6, Beh (2004).

8.3.4 The second OASIS assignment

Prior to this assignment, all assessments had to be set up by the OASIS programmer himself. Recent software enhancements now made it possible for academics to set up their own assessments and this was the first such assessment. A screen-shot of the OASIS assessment setup page is shown in Appendix five, Figure A-5-9. After specifying the assessment as an assignment rather than a test, I readily specified the questions in the assessment, the start and finish times and dates for the assessment, and its duration. In this case it was to be done by each student in a one-hour period

between 8 am and 8 pm. I also named the assessment and placed instructions for students on its first page. The assignment itself took place on the Friday of week eleven of the twelve-week semester.

As outlined in Subsection 8.3.2, there had been major problems with the first year-one assignment and a few students had expressed a lack of confidence in the ability of OASIS to deliver assessments smoothly. It was essential for the credibility of OASIS and the course that this second assignment ran without problems. Pleasingly, this was the case. In particular, the server handled the load extremely well, its full capacity never being approached.

Student performance in this assignment was also encouraging. The average mark in the first assignment had been 75%, with 125 students gaining full marks and a further 89 students gaining 12 out of 13. For the second assignment the corresponding figures improved to 88%, 279 and 108. While these figures did not quite equal the corresponding figures for the second year-one assignment of 2003 (92%, 310, 100), at least they were closer to them than the first assignment's figures had been to the corresponding figures for the first 2003 assignment. Given the fact that the 2004 assignment had relied mainly on previously unseen questions while the 2003 assignment had been based completely on previously-seen questions, this difference in assessment performance was perhaps to be expected. It certainly was not evidence that the class of 2004 grasped the material less well than their 2003 counterparts. In fact, I was pleased that I had been able to make the transition to previously-unseen assignment questions without incurring a large drop in the students' marks.

8.3.5 Student voices: interviews and emails

As outlined in Subsection 5.7.2, students were invited by email to take part in interviews. Ten students subsequently volunteered to be interviewed. In these interviews I was particularly keen to explore the issue of the previously-unseen assignment questions, since that was the one new issue this semester. Much student feedback had already been received about other aspects of OASIS and its implementation; feedback on these aspects was expected to be confirming rather than groundbreaking.

In response to my emailed invitation, a number of students also replied with their thoughts on OASIS. Some quotes from these student emails are also included in this subsection. Such quotes are noted as being from emails. The majority of quotes are from interviews and are not explicitly identified.

I had hoped that the assignment's previously-unseen questions would effectively eliminate the use of spreadsheets. This appeared to be the case: only one student voiced any concerns about their use, and that student appeared not to realise that the assignment questions no longer were chosen from the practice questions:

> S: I think I mentioned in my email how there were people selling spreadsheets as well, the night before, even third-year students trying to sell spreadsheets and that's a very common sight. I can understand why that happens but still….
>
> CS: The problems used in the assignments weren't identical to the ones used in practice so….
>
> S: Were not identical?
>
> CS: That's right, so if anyone had made a spreadsheet two years ago, or even the night before, it wouldn't be very useful...

However, another interviewee did comment that some of his peers were still hoping that spreadsheets would be useful, in spite of having been told that most of the assignment questions would be ones not seen before:

> Previous groups had suggested to other first-years to actually make spreadsheets... And so therefore a lot of them were actually expecting going into the assignment that it was going to be relatively the same and it was just a matter of finding which questions they were so I think quite a few people would've found that quite shocking.

As in the past, there certainly was considerable group-work during assignments, and the general opinion was that this gave group members a distinct advantage:

> CS: Can you please tell us how you found OASIS this year?
>
> S: I found it was really good as far as learning the work and practising the work.
>
> CS: Okay. Yes.
>
> S: I thought it was quite a bad way to test everyone because it kind of seemed like it was pretty much testing how well you could collaborate with everybody else.
>
> CS: So you thought there was a lot of collaboration involved?

S: Oh definitely. The first time I did really badly because I did it by myself and everyone who did it badly, did it by themselves and everyone who did well seems to have done it in groups.

However, since each student in such a group had their own version of each assignment problem, group-work did entail doing each problem more than once:

> When we do OASIS we actually do it in a team. Like three or four people. So after we actually practise all the OASIS practices, so it's kind of like we just spend fifteen minutes for our assignment, so another forty minutes we check each other's assignments so that we make sure that our team, like our race, get high marks.
>
> We did the practices before so when we do the actual assignment it only took us about 20 minutes. So we actually have 40 minutes more. As for me I have two team members so I spend 20 minutes checking theirs. I redo everything. I spend 20 minutes and I redo everything. I spend another 20 minutes on another person. So they did the same too, like there's three of us in a team, so we help each other out.

Some groups appeared quite well organized, with the members having written methods for each practice question. However, with 560 students sitting an online assessment in a 12-hour period, and with many of those sitting it at the University, even students who had not set out to organize groups might find themselves in a room with several others also sitting the assignment. In such situations informal assistance did often take place, bringing people "closer together" and producing a "stimulating study environment":

> For OASIS assignment it's like you do it in an environment which everyone is just doing it and helping each other and you can actually ask questions even to other people who are doing it and you don't know them. Like everyone just gets closer together and they just discuss together and that environment is, I think it's a better stimulating study environment rather than just a test.

As the above suggests, students did not always view assignment group-work and collaboration negatively.

> The only thing I thought a bit flawed about OASIS was that it was easy for friends to get together and help each other out. I am guilty of this, but I also found this group-work was ideal for learning. So ultimately, I found OASIS was a real help for my learning. [student email]

Taking the above further, one student noted that OASIS encouraged team spirit:

> Engineering is actually about team spirit and about team work... OASIS indirectly it actually instils something about team spirit. They actually instil in everybody the team spirit, like how important the team spirit is. Like very few people actually do it on their own, the assignment. At least we do it with more than one friend. Yeah. Like we can see on Friday where the OASIS

> assignment, everyone just, you know, a bunch of people had moved together. Yeah. With that way you learn in, yeah and you form a team spirit. Yeah, which is really good.

Interviewees also commented on the fact that some students would wait till others had done the assignment, or would watch others doing the assignment, in order to find out the questions in the assignment. These students would then practise those questions prior to sitting the assignment themselves:

> Whoever sits it first is at a disadvantage to their marks because they're the first one in.
>
> The people that do it first are normally the guinea pigs. You know what I mean and they can write out how each problem was done and people who have had no practice before, they don't understand anything, they can just go ahead and score full marks.
>
> There are a lot of opportunities to cheat just because the test questions are the same. So whoever does it first, quite often what would happen is they would just print out the questions, and for the rest of the day people would just look at the questions to get the answers and they just got 100%. That was a bit unfair on the ones who were doing it properly and didn't get 100%, when they just used the questions and got 100%. And they didn't even do the OASIS questions in the first place.

This had been less of a problem previously: when assignments were based solely on previously-seen questions, students could prepare for the assignments by simply making sure that they could do all the practice questions. Often this did entail writing out how to do each question or even producing a spreadsheet. With the advent of previously-unseen questions, students could no longer rely on these methods, and it seemed that the two other approaches described above (group-work and determining the assignment questions prior to sitting the assignment) were becoming more prevalent as a result. Although students had been told that the unseen questions would be very similar to the previously-seen questions, not all students were confident enough to rely on their ability to solve these previously-unseen questions, and this no doubt increased the tendency to seek help:

> Not many people are brave enough to go first. I certainly wasn't, I just waited until someone else went first, asked him or her how he or she did…

Further, some students who had insufficient understanding did not perceive the unseen questions as being similar to the questions they had practised, further increasing the problem:

> Yeah, it's a good idea to make them similar, just like, actually similar to the questions, which a few of them weren't. I noticed that, especially with the assignment questions.
>
> I found the difference between the questions that were practised on and the questions that were in the test for the first time, they were quite different and I didn't feel like I'd fully got enough understanding of the different circuits until this round of studying...

Others were not troubled by the differences between the assignment questions and the practice questions:

> If you worked through all of the practice questions, you're pretty good for the test.
>
> CS: The problems in the assignment were a little bit different to the practice problems; did you find that a big difficulty or not?
>
> S: Okay. Not really because the most important thing is you have to understand the concept, you see.
>
> I liked the way it had just a certain number of questions... it wasn't so broad that you didn't know what to start with and that sort of thing, so I liked that. And I liked the way, because you knew it was going to be in the test, you did do all of them. Well, I did all of them, so it gave me an opportunity to actually get good rewards for what I studied... I do this and if I know it then I will get good marks.

Much of what students said during interviews paralleled earlier student feedback. Again there was no clear-cut majority for either tests or assignments. Those that preferred tests did so because assignments allowed students to cheat and therefore could not accurately measure student achievement:

> Assignments do not represent students' capabilities since large amounts of cheating activities are present:
> (a) students helping each other
> (b) students work in large groups
> (c) students looking at others doing their assessment and copying all questions down before trying their own
> (d) students printing all practice and looking at related question during assessments so they have previously-worked solution in front of them
> (e) students who have not learnt the theory at all ask so-called good students to check their answers if not do the assessment for them completely
>
> [student email]
>
> I believe that OASIS is an extremely useful tool for building up concept familiarity and for exam preparation, although not particularly effective nor accurate as a means of assessment. Just walk into the level-three computer labs on the day of an assessment and you'll see hordes of first-year losers (myself included) crowded around a single computer, busily copying the test questions, memorising the solutions, giving/receiving/demanding excessive help, etc. People selling spreadsheets they had made the night before, covering each

> question, are also a common sight. Your performance is dependent not on your amount of knowledge or practice, but on the number of people you have beside you or how rich you are. [student email]

Those that preferred assignments mentioned the learning that went with group-work and the relative lack of test anxiety:

> I think OASIS assignments are better... because you get really stressful when they say it's a test, you see. You just get paranoid about the word "test" because it's like you are being supervised and yeah, in such a stressful environment, like everyone sitting quietly
>
> I think it's a better stimulating study environment rather than just a test.
>
> Because there's no pressure I found I did actually pay attention to what I was learning and I learned more from that than what I usually would because I was paying attention and I wasn't like 'oh quickly, quickly get all this work done so I can get good marks'.

The interviewees often spoke favourably about the way OASIS had helped them increase their understanding:

> As far as a study tool I think it's really good. I find it really helpful. I've used it heaps for studying for the exams. Yeah, and it's just a help and that's more like before you originally understood it, just not really getting what's going on.
>
> I thought the OASIS was very helpful this semester. I would like to point out that OASIS helped me in understanding the concepts behind each problem more clearly. I agree that it is a great learning tool in the course and would consider it more worthwhile to keep it going. For me, it has definitely helped to develop a stronger understanding in electrical circuits, induction and electrical motors etc... Thank you for your knowledge this semester. [student email]
>
> I reckon it was really good though. Like, I phoned a few friends to tell them that I was coming to talk about it and they said, "Yeah, it's an awesome exam resource. It was really good". I actually learnt how to do the problems. Like, there's different levels. You can know how to do it and you can understand why you're doing it, and then you can just do it automatically, if you know what I mean? And it took it from that level of knowing how to do it to understanding why you do the things that you do.

In particular, the fact that questions could be repeated with different numerical values was seen as helping students improve their understanding.

> Every time the values change you really cannot actually directly get the answer from other people so in that case it means that the only way you can get your correct answer by your own values. It can only be really learning and realising what's the correct way. What correct principles could be used.
>
> I've probably gone through all of the questions on the circuits about five times completely and the questions I don't understand I've gone through more times. Because I feel like I don't understand it and I need to retry it with different numbers three or four times just to make sure I know what's going on.

> The best part about OASIS compared to problems taken from books is definitely that you can practise the same question over and over but the numbers change each time, so when you do the problem again and get it right, you know it's because you actually understand how to do it as opposed to that you remember the right sequence of numbers from the last time you did it.
> [student email]

However, it was also noted that some students appeared to adopt a formulaic rather than an understanding approach to their problem-solving:

> I found when I asked people how to do a particular question that I didn't understand, they would get the right answer but it was just from seeing how to plug in numbers, it wasn't so much from understanding the theory behind it. I'd ask someone 'how did you do question 16' and they'd say 'oh, yep, there, there, there, that answers this' and during the practice, you'd test it out and it was right and I'd ask so what happened and they weren't too sure themselves. But I guess there's only so much you can do about that.

Such an approach could have been adopted by students who believed they did not have sufficient time to develop a good understanding of the material. Certainly, students did comment that time-pressures often affected their study patterns:

> The second test, I just did them all once. The first one, I did them all until I got them right and I could do them on my own without looking back and having to check what the formula was or something. The second one, I didn't have much time so I just did them all on the last day.
>
> Engineering as you know is really stressful, there is always something that is more important than something else, like there would be Design which is due on the Wednesday, then OASIS on the Friday, so you work on Design until the Wednesday and then at least you've got those days to focus on OASIS.
>
> The only problem is the timing of the OASIS assignment. The first one was before the mid-semester break, so we had to rush our study and instead of understanding the concepts, we just guessed our way to find a model answer for each question. Therefore, if you were to have it after the break, we could have time in the holidays to go over the material covered and then answer the questions with our knowledge, rather than our guess work. [student email]

In general, the interviews confirmed previous findings. They also highlighted some issues with assignments: whilst the use of previously-unseen questions had eliminated spreadsheets, there was a possibility that it had increased the incidence of other approaches, such as working in groups or finding out the assignment questions from 'guinea-pig' students.

8.3.6 The final examination

Students gained an average mark of 81% for the OASIS assignments. This did not reach the standard (87%) set in 2003, but I judged the difference to be attributable to problems with the delivery of the first assignment and the change to previously-unseen assignment questions. I did not consider that there had been a significant drop in achievement. The average mark for the two tests was 58% and, for the final examination, 60%. The corresponding figures for 2003 had been similar at 62% and 58%, respectively.

The final examination followed the format of the 2003 examination. I wrote two of the six questions. Just these two were supported by OASIS questions. They returned averages of 71% and 65%. The other four questions averaged 56%. The corresponding figures for 2003 had been 54%, 68% and 56%. With the exception of the result for the first of my questions, student performance in 2004 was very similar to that in 2003. The large increase in the average for that question, I considered, was a result of the 2003 examination having identified some student weaknesses; in 2004 I specifically targeted and largely corrected these weaknesses through writing further OASIS practice questions and through reducing the course content marginally and using the time gained to promote deeper understanding. These events were foreshadowed in Subsection 7.6.4.

8.4 Reflections and plans for future action

The new version of OASIS replaced the prototype version in semester two, 2003. Thus 2004 was the first full year of OASIS usage, and it was gaining more recognition as a mainstream teaching and learning tool. Use of OASIS continued to grow within the Department of Electrical & Computer Engineering. The Department of Physics had also adopted it enthusiastically and was using it for classes as large as 470 students. Other departments were showing an interest in the software too. There was a high usage of OASIS in courses where it was offered, with both students and staff speaking favourably of the impact on student learning. Again, student interviews, emails and surveys corroborated informal feedback.

With more large classes using OASIS, assessment scheduling became more important. Records showed that OASIS usage was heavier during the last day of practice for an assessment than for the assessment itself. The Department of Physics always consulted us about assessment scheduling and we never held more than one assessment on the same day. However, sometimes assessments were on consecutive days, generating particularly heavy usage on the day of the first assessment, when one group would be taking an assessment while the other was practising for their assessment the next day.

At the start of 2004, two year-four students had embarked on a research programme which aimed to maximise the efficiency of the OASIS software. This was certainly most timely, for a number of reasons. First, the software appeared to be running more slowly as more features were added to it. Second, there were now more users of the software, in particular the Department of Physics. Third, this Department often used questions that contained relatively large graphics that made greater demands on the server. The server load for August 4, the date of a large-class Physics assignment, is shown in Figure 8.4. It can be seen that from 8 pm to midnight the load on the server's CPU was approaching the maximum permissible load.

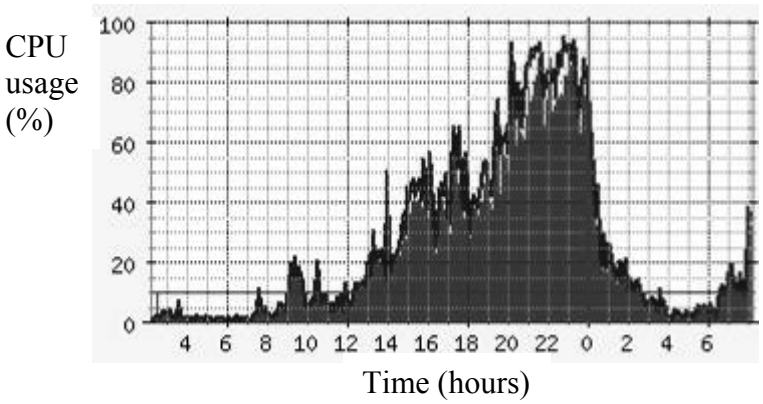


Figure 8.4: CPU usage against time for August 4, 2004. The CPU usage is expressed as a percentage of the maximum load. The period of time actually shown runs from 2 am August 4 to 8 am August 5.

Unfortunately, soon after, on August 27, during a year-one Electrical Engineering Systems assignment, the system failed to cope with the load. This failure may not have been entirely attributable to the OASIS system or software: the university-wide Cecil system completely failed at the same time. However the affected students

typically did attribute the failure to OASIS. Consequently it became imperative to make the OASIS system as reliable and capable as possible.

After the above failure, various performance restrictors were identified. One restrictor was the time taken to retrieve images from the data-base. This was addressed through the reduction of image sizes where possible, together with the caching of static images. Another restrictor was the requirement that all variations of each question were being retrieved from the data-base when only one variation was needed. This restrictor was addressed through a software upgrade. The server RAM was also tripled in size. Once these changes were implemented, it was considered that the system was easily capable of coping with the current loads as well as any loads expected in the near future.

Given that OASIS was now regarded as a mainstream teaching and learning tool, it was important to future-proof it against other eventualities too. For example, there was just one programmer with the knowledge and expertise to fix software problems. This was unacceptable for software extensively used by several courses. To make the software less reliant on one person, the programmer documented the software. A postgraduate student was also employed to document the software independently. Another postgraduate student, guided by a member of the OASIS Committee, was employed to write a comprehensive user-manual. These documents were completed by the end of 2004. Given this documentation, and the software and hardware upgrades noted earlier, it was judged that academics could confidently make a long-term commitment to using OASIS in a wide range of courses.

The University's Department of Physics had successfully trialled OASIS in 2003. In 2004 the Department adopted it as a mainstream tool in courses that involved several hundred students in total. When interviewed, the academic responsible for this implementation commented that, by allowing students to do the same problem again and again, OASIS helped the students to gain a better understanding of the physical models involved. He also stated: "We find OASIS particularly useful for first-year students as the online learning system means they are able to learn in their own time and at their own pace". Physics assignments were similar to electrical engineering assignments except that the former were available over a 24-hour rather than a 12-

hour period. Automatic marking was seen as a huge advantage, saving considerable staff time. These assignments were not an addition to the assessment regime; rather they replaced traditional in-class open-book assignments. These assignments had been marked overnight by teams of markers who began the marking at 7 pm in the evening and who undertook to complete the task and provide the marks to students by the start of the next university day. This operation was unpopular with markers, involved considerable marker time and effort, and raised questions about marking accuracy and consistency. By contrast, OASIS was able to deliver automated instant feedback for very little staff effort. Additionally, OASIS provided statistical data to staff, enabling them to identify and act on weaknesses in student understanding:

> With OASIS we not only get average marks for the test, but also average marks for each question, which allows us to pinpoint concepts which are difficult for the students. The system also allows us to find out how much time students spend on each question as well as the time of day that they took the assessment.

It was also noted by the academic interviewed that, while OASIS was currently being used only at university level, it was highly applicable to high-school physics.

Another sign that OASIS was becoming more accepted was the commercial interest now being shown in it. During 2004 the OASIS committee had had meetings with UniServices, UoA's commercial arm, and with representatives from Cecil, the University's learning management system. Although the committee members were not united in their support for linking OASIS closely with Cecil (some regarded Cecil as too slow and cumbersome), it was decided by majority vote to explore the possibility of providing OASIS via Cecil. Thus, students who were logged on to Cecil could seamlessly access OASIS for both practice and assessment. OASIS assessment results would be stored automatically by Cecil. UniServices was keen to provide OASIS within Cecil because it saw Cecil as having a commercial value that would be enhanced considerably by the addition of OASIS. In order to investigate the feasibility of the linkages, two students were employed over the summer of 2004-5 to determine how compatible the software packages were. It was found that considerable modifications would be needed to enable the two packages to work together. Consequently, the decision was made, midway through 2005, not to pursue the idea of providing OASIS within Cecil.

As outlined above, in 2004 much effort went into consolidating and documenting the progress of previous years. These efforts produced a more stable, robust and mature operation. Five important advances were also made in the features offered by the software itself. First, instructors could now monitor student progress live during assessments. This proved to be particularly useful when students had problems with assessments, such as a crashed computer or a submission accidentally made before assessment completion. Second, a student look-up facility enabled academics to determine the details of practice and assessment activity for any student. Similarly, other software improvements made it easier for staff to access data about class performance on particular practice questions or complete assessments. Third, the ability to set up questions with consequential marking enabled students to receive credit for partially correct working. Fourth, academics could now set up their own assignments; previously this required the intervention of the OASIS programmer. Screenshots illustrating some of the above functions may be found in Appendix five. Fifth, in 2004, previously-unseen questions were used in assessments for the first time. From the student perspective, this was the only significant change for 2004. Of course, the 2004 year-one students would not have been aware of the change.

There were two reasons for the introduction of previously-unseen questions in assessments. First, it was considered that such questions would lead to assessments that were fairer and perceived as fairer. I had been told that some students used spreadsheets prepared by others to help them with assignments. This problem seemed to exist much more in the year-one course than in the year-two course. Several year-one students had complained that spreadsheet use during assignments was unfair. Subsection 8.3.5 presented and discussed student comments on this matter. Second, it was hoped that the introduction of previously-unseen questions in assessments would dispel doubts that impressive results in OASIS assessments were a product of rote-learning of the practice questions. In fact, the students' results in traditional paper-and-pencil tests and examinations already provided clear evidence of genuine learning. It was also noted that the introduction of previously-unseen questions did not significantly lower student performance in the assignments.

The introduction of previously-unseen questions in assessments did virtually eliminate spreadsheet use. The assignments still provided good motivation for the

students to practise the OASIS questions: most saw the practice questions as similar to the assignment questions. Students had made the point that it was important, for motivational purposes, to preserve the close relationship between the practice questions and the assessment questions. The consensus was that the relationship was about right. The fact that the practice questions were not identical to the assessment questions encouraged the students to think more generally about their problem-solving skills, rather than to concentrate on being able to solve particular questions.

However, the introduction of previously-unseen questions in assessments had not eliminated questionable practises. While students no longer used spreadsheets, it was clear that two practices had been adopted by students who didn't want to rely entirely on their own abilities. First, some students observed others doing their assignments, recorded the details of the questions, practised them till they felt confident they could solve them, and then sat the assignment themselves. Second, some students worked in teams on the assignments. In some cases this involved the students in extra work – each student in a group typically doing each question three times – but it was seen as producing higher marks.

In general students were very positive about the way OASIS had improved their skills and understanding, but they were less happy about its use for assessment: OASIS assignments were regarded by some as unfair because they were open to cheating. I wanted to reduce the opportunities for cheating as much as possible. Although our practices were already relatively strong in this area, with the use of numerically different versions of the same questions, the use of different images within one question, and the use of previously-unseen questions in assessments, I judged that still more was needed. The problem could have been eliminated by moving to supervised tests. However, I considered this to be impractical. First, whereas the students were reasonably accepting of regular assignments, it was clear from interviews that they would regard a series of tests with their associated high stress levels very unfavourably. Second, additional tests would require considerable organisation in terms of rooms and invigilation. Further, it was the year-one class that had voiced the most negative view of cheating during OASIS assignments and this class was too large for supervised CBA: we did not have sufficient computers that could feasibly be used for a supervised test.

To render assignments less open to cheating, I decided that future assignments should provide different students with different questions, not just numerically different versions of the same questions. I judged that, in my two courses at least, there were sufficient OASIS questions to enable each assignment question to be chosen from a small number of alternative questions. This improvement required further programming which was completed in time for the 2005 academic year.

I also wanted academic staff to be able to access a wider range of statistical data on student activity, and to do so more easily than was currently the case. We had previously hired a postgraduate student to work in this area over the summer of 2003-4. However, more work was still needed, and another student was hired to carry this out over the summer of 2004-5. In particular, I was keen to gain an insight into the time students typically spent on each OASIS question prior to submitting their answers. I was also interested in how long students were spending on their OASIS assignments. Over the following summer some good progress was made in these areas, in spite of a few delays: in December the entire Department shifted from one building to another, and in February the OASIS development server failed and its replacement also failed. Considerable functionality of a simple statistical nature was added to OASIS, and good improvements to accessibility and presentation were also made. Much information about student performance could now be presented graphically. One unexpected problem was also identified. While the database did have the functionality to record the times students spent on questions, it was not actually recording these times. This problem was fixed by the start of the 2005 year, enabling such data to be collected during that year.

In spite of the system failure during the first year-one OASIS assignment, I considered that, overall, the two action-research cycles of 2004 had been most successful. In terms of outcome validity, there was a good body of evidence from examination results, student surveys, and staff and student interviews to suggest that student learning had been enhanced. This evidence was consistent with that from the previous year. The evidence from the survey had been so consistent across three research cycles that I judged it unnecessary to continue using it The evidence of student learning provided by assignments in 2004 was somewhat more reliable than

that from 2003, given the use of previously-unseen questions in 2004. From my personal perspective, the desired outcome of staff workload reduction had not been achieved. Much of my workload I judged to be attributable to the fact that we were still improving OASIS itself and its implementation. My role as Convenor of the OASIS Committee also added somewhat to my workload. I still hoped that, once reasonable stability had been reached, I would enjoy a reduced workload. The Department of Physics certainly had enjoyed a reduced workload after implementing OASIS for assignments in its large year-one classes.

In terms of process validity, I judged the rigour of the research to be high, on a par with that of 2003. In addition to valuable informal feedback from students, formal feedback from a dedicated survey and taped and transcribed interviews was again collected and analysed. The survey forms were, as previously, analysed by CPD. Triangulation was further increased by the involvement of the University's Physics Department: they had moved from a trial of OASIS in 2003 to an enthusiastic adoption in 2004. It was clear from their comments and actions that they considered OASIS both promoted student learning and reduced staff workload. Their involvement provided further observation, method and interpretation triangulation.

The democratic validity of the action research was maintained through informal channels of communication with staff and students as well as the more formal channels provided by interviews, survey forms and the OASIS Committee. The growing Department of Physics commitment to OASIS also increased the numbers of students and staff that were exerting some influence on the research programme.

The action-research cycles planned for 2005 would collect further data from assessments, interviews and the like, thus providing the opportunity to increase the validity of the research. I planned to record and transcribe interviews with relevant academics as well as students. Further software development was again planned to enable better access to the quantitative data collected and stored by OASIS, thus generating further possibilities for validating the research.

8.5 Summary

There had been considerable change in 2003, most notably the migration from the OASIS prototype. By contrast, 2004 was a year of stability, with much of the activity centred on securing OASIS as a mainstream learning and assessment package capable of handling more than one large class concurrently. The server's RAM was tripled and software performance testing led to upgrades that enhanced the software's ability to handle large numbers of users. Extensive software documentation and a user-manual were completed to provide some protection against staff changes. Consequential marking and a student look-up facility for staff were also implemented. Academic staff were now able to set up assignments and monitor student progress during them.

For students, the main change was previously-unseen questions in assignments. This change did not reduce assignment marks significantly. Interviewed students cautioned that, to motivate practice, the assignment and practice questions should remain similar. There was also debate in student discussion groups about hints; some students saw them as most helpful, others as an unrealistic crutch. The decision was made to continue not to provide hints.

Previously-unseen questions were introduced to counteract spreadsheet use in assignments. However, students now reported other forms of assistance: assignments sat by students working in small teams, and students learning the assignment questions from others taking the assignment before them. These students did not view OASIS as a fair assessment tool, though they did view it as valuable for learning and made good use of it. To reduce the above problems I decided that future assignments, particularly at year-one level, should offer different questions to different students, not just numerically different versions of the same questions.

The Department of Physics adopted OASIS as a mainstream learning and assessment tool in 2004, and their own perspective added a good measure of validity to my research. Their input of ideas and financial support for the development of the software was appreciated too. This state-of-affairs continued through 2005, the subject of the next chapter.

Chapter 9. Spreading the research net wider: 2005

9.1 Purpose and outline of chapter

The previous chapter outlined the events of 2004, the first full year of the implementation of the new version of OASIS. This chapter describes the events of 2005; again covering two semesters and therefore two further cycles of the action research. The year-two course 'Circuits & Systems' is the focus of the first cycle, described in Section 9.2. The second action-research cycle, involving the year-one course 'Electrical Engineering Systems', is discussed in Section 9.3. No large changes were implemented in these cycles. With the software now undergoing few changes, the main focus of the research was on the second and third research goals (presented in Subsections 5.6.2 and 5.6.3, respectively). To improve triangulation in addressing the research questions associated with these goals, evidence was sought from three further sources. First, the implementation of OASIS in the Department of Physics, together with the staff and student response, was investigated; this is the subject of Section 9.4. Second, it was deemed important to explore the longer-term consequences of the implementation of OASIS. To this end, postgraduates who had, as undergraduates, used OASIS were approached for feedback. Their feedback is presented and discussed in Section 9.5. Third, I interviewed the ECE academics involved in OASIS, these interviews being the subject of Section 9.6. No formal interviews had been conducted with academics since 2003 (see Section 7.8). This second round of staff interviews served two purposes: first, to provide another longer-term perspective that might confirm that of the postgraduates, and, second, to update and confirm the staff perspective obtained in 2003 by a different interviewer on an earlier version of OASIS. Finally, reflections and plans for future action are outlined in Section 9.7, and the chapter is summarized in Section 9.8.

9.2 Year-two, semester one, 2005

This semester I was again involved in teaching the year-two course 'Circuits & Systems'. The course and the lecturers were both unchanged from the previous year. I considered that my part of this course had run very well last year and I implemented no changes for 2005.

9.2.1 The OASIS assignments

As previously, there were four OASIS assignments and I was responsible for the first two, each worth 2% of the final grade. Again, the first assignment took place on Wednesday of the second week of the semester while the second assignment took place on Wednesday of the fourth week. At this time, OASIS did not have the ability to supply different assignment questions to different students. However, as in 2004, previously unseen assignment questions were used. Both assignments comprised six questions. The first assignment had three previously-unseen questions, the second, five. The assignments returned average marks of 89% and 86% respectively, consistent with previous results. Both assignments ran very smoothly, with few problems reported.

The ability to monitor the progress of students as they completed their assignments produced some interesting information. For example, it was clear that most students did not unduly delay assignment submission. For both assignments, approximately 85% of the students submitted their answers in the first half of the available twelve-hour period. It was also gratifying that, in general, students who submitted their assignments earlier scored more highly than those who submitted later. Had there been a major problem with students observing their colleagues taking the assignment and then practising the assignment questions, the reverse trend might have been expected. The same pattern of earlier submissions scoring more highly than later submissions was also observed in the Department of Physics.

The newly-available data concerning the time spent by students on their assignments also provided some valuable insights. For example, the year-two students spent, on average, about 45 minutes on their assignments, a not-unexpected figure. However, a few students completed the second assignment very rapidly. For example, during the second assignment, shortly after 9:00 am, one student scored 100% in only 11 minutes, a remarkably brief period of time. This particular student had scored a B+ in the year-one course and 90% in the first assignment in the year-two course, taking 31 minutes to do so. The score of 100% after only 11 minutes I therefore judged to be somewhat out of character. However, only five students had completed the assignment before he had, and the assignment had been available for such a short time that it was not possible for him to have practised and mastered the questions

prior to submitting his answers. This surprising result was overshadowed by a few more as the day progressed, most notably a mark of 100% scored after only six minutes. I didn't have any completely satisfactory explanations for these results.

9.2.2 Focus-group discussion

I had been very pleased with the quality of information gleaned from the year-two focus-group discussions in 2004 (see Subsection 8.2.3). Therefore I decided to conduct a focus-group discussion again in 2005 in preference to individual interviews with students. The discussion took place in my office, involved four students, and was about 40 minutes long. The information gathered was confirming rather than ground-breaking. In general, the students were positive about OASIS, stating that it did help them learn. They also spoke of the motivating power of assessments and how OASIS was useful preparation for traditional assessments as well as the online assignments:

> It's very good, a good exercise and good practice for the test and of course for the assignment.
>
> The work keeps you practising constantly because we have tests, assignments all coming up.
>
> To me I think I need to do the practice when there's an assessment task coming up. If you give more tests and assessment I think it helps for me to keep on practising because I know there are tests coming...
>
> I think OASIS is pretty good. I don't buy text books… I need to practise before the test, if I can't find anything else, I can go to the Internet to do the OASIS practice again. For the other papers they don't give us much practice, for us if we practice then we will understand, then we can do even harder questions but for the other papers they only give us the notes… they don't give us some extra practice… but for this paper you give us OASIS and we can practise…

However, students did also comment that it was possible to get some questions right without a proper understanding and that this state of affairs could persist because the software marked only the answer and not the method. There were also, as there had been in the past, some requests for explanations as well as answers.

> I was wondering if there was an explanation provided why we got the answer incorrect... this is the method we use just because we got the answer like that but not actually understanding, we know that we use the method like this and we got the answer and next time I see a question like that I'll actually use the same method but without understanding the actual concept.
>
> I think OASIS will be much better if it provides methods of solving the questions... if you try to solve a question with a method, that it can actually check your methods... we learn more from that.

One student noted that, in his experience, it was actually possible to use an incorrect method and repeatedly get the right answer for some questions:

> S: My concept was totally wrong but I managed to get the answer... I assumed stuff in the first place already which was wrong so sometimes the concept, explanation would be good on the [OASIS] site to tell you what you actually did wrong.
>
> CS: If you had repeated that question... would that have shown you that you were doing it the wrong way or would your wrong way always have worked?
>
> S: Yeah always would have worked but the concept the way you think is not the correct way... Most of the time it will still work, just the concept is not correct.

Following the above discussion, the suggestion was made that OASIS could mark not just a final answer to a question but could also mark intermediate steps on the way to the final answer:

> S1: This one question, you divide it up into small sections and you come to input answers for each small working, then if you get one wrong you know whereabouts you got it wrong and then you know what concept you got wrong and you go back and have a look, rather than just putting the final answer...
>
> S2: That's true; I was thinking that as well... if you've got a complex circuit like Thevenin equivalent, overall the only two questions that we can answer in that case are Thevenin resistance or Thevenin voltage but it would be good if you can just ask us a question in the middle... just like you said like a question should be divided into small parts.

As I pointed out after the above comments had been made, asking for intermediate answers presupposes that a particular solution method is to be followed for the question. However, there was a suggested solution for that problem too, albeit a somewhat unwieldy one:

> How about instead of just providing answers at the end... you provide solutions with it, like nodal analysis, this method, and another method which is mesh-current and students can check which?

The textbook was also discussed. As had been found previously, OASIS questions were done in preference to questions set from the text, even though students did have answers for all set text questions. When the text was consulted, it was usually in an attempt to understand the theory or to solve an OASIS question:

> If we don't understand we can refer to the textbook from time to time and just get out the concepts from there but actual practice we are doing it in the OASIS

because we try the question about two or three times till we have got the concept in our head.

Again, there was confirmation that some OASIS practice was done offline. Apart from the fact that Internet access was not always available, there were two main reasons for this, the first being the ability to add one's own notes to a printed version of the OASIS questions, and the second being the eye-strain associated with extended session on some computer screens:

S1: What I did was copy and paste all the questions into Word and then print it out and also when I did the questions I just wrote the solutions on the printed pages so like in the end like before exam or before the test I can just look at the pages I printed out and also the solution I already did.

S2: Sometimes it gets quite tiring after two hours on the screen; your eyes do get quite sore.

CS: Were you doing them always online or did you also print them out?

S2: I do online but that would take quite a while because I copy the stuff but sometimes the picture won't pop up, sometimes I just get a blank, but what I do I just go on one hour and take half an hour break and then other times go up, go more frequent rather than go a long time.

The students compared the coverage of the assignments, noting that the first one covered relatively little material, while the second covered too much. In order to solve this problem, it was suggested that some content be shifted from the second assignment, either to the first or to a new assignment. Perhaps surprisingly, there was no objection to the creation of an additional assignment. In fact one student suggested that there should be an OASIS assignment fortnightly or even weekly:

S1: the second assignment that we did I remember we had heaps of questions to do: the AC and the DC and the Thevenin. Our group, we said, I wish he had separated this assignment into two parts... we would have liked that. It was like so many questions and only five were coming up but at the end of the day when the questions came up not everything was tested, there were no question on attenuators...

CS: Would it have been better if I had moved some of the material in the second assignment into the first assignment perhaps?

S2: Yeah. The first one was quite easy.

S1: Or had another one... More like on a fortnightly or weekly basis probably would be better.

This discussion took place two weeks after the third assignment in this course. This assignment, from another lecturer, could and did make use of randomly selected questions. The specification of this assignment was very simple: it comprised just

one question chosen randomly from a set of seven previously-seen questions. Some students had been surprised that different students received different questions, although this innovation had been communicated to them. Generally the assignment had not motivated a great amount of practice since there were only seven questions and, once students had worked out how to do these, they considered that no more preparation was necessary. It was also noted that a one-question assignment could end in disaster for students who made just one mistake. On this basis an assignment consisting of more questions was seen as fairer:

S1: The third one [assignment] was quite different because of different questions. She [my friend] told me it's just a normal question and then after that I went to do it and I said 'oh, my God'.

S2: It changed quite a lot.

CS: So you each got one question and you got different questions. How do you feel about that?

S3: Quite interesting, but one question is sort of a bit risky, right? If you're wrong, that's it. You fail. So probably three or four you'd get it rather than just one.

CS: All right, so if there were, say, six questions and you knew the assignment was going to be three out of the six, you're saying that might be a bit better? Would it be? Because you're saying that if there's just one, it's a bit risky because if you make one careless mistake it's gone.

S1: From the student's point of view I should say. Before this assignment I was really prepared for the first two assignments because I knew any six [questions] could come out of those and there were so many to do but the third one when it was about to come I wasn't that prepared. I wasn't that motivated to do it because I said 'only one of these is coming and we have got all the solutions written out'… And then I saw two of my friends and they said 'oh, we had question number one' and then when I went to do it I got question number three and I'm like 'oh, my God'! I knew how to do it but I wasn't that motivated to do it, that's for sure.

At the end of the discussion I turned off the recorder. After this, the students made more relevant comments. I attempted to write them down after the students left the room. The most notable comments were to the effect that, for an upcoming test, the topics areas for which there were OASIS problems needed far less revision because those areas were much better grasped than the other areas: "we don't need to revise the OASIS parts so much; we just concentrate on the other parts".

9.3 Year-one, semester two, 2005

9.3.1 Changes from 2003

There had been few changes in the year-one course 'Electrical Engineering Systems' in going from 2003 to 2004. However, there were two significant changes in going from 2004 to 2005. First, one lecturer for the course was on leave, and rather than find a person to deliver his set of five lectures, it was decided to dispense with them altogether. While this reduced the course content by about 15%, such a reduction would be necessary for 2006 anyway when the University would be restructuring all its degree programmes, a move that would require the course to undergo just such a size reduction. It was decided that a reduction coming one year early would be better than finding a replacement lecturer for just one year. The consequence of this and some minor changes was that I gave half the course's lectures.

Second, it had been decided at Faculty level to accept 600 students into first-year Engineering, a significant increase over the 560 accepted in 2004. This increase was likely to produce a poorer-performing student body than had been met previously. Further, the students starting university in 2005 had experienced the first three years of a new national assessment regime in high school, the National Certificate of Educational Achievement (NCEA). While this should not have, in itself, led to reduced standards, the fact was that these students had experienced the first year of year-11 NCEA in 2002, the first year of year-12 NCEA in 2003 and the first year of year-13 NCEA in 2004. I believed that, with their teachers preoccupied with adjusting to and implementing the new system, these students would be less well-prepared than previous students entering the Engineering Faculty.

The assessment of the course was identical to that of 2003 and 2004, described in Sections 7.6 and 8.3, respectively. As in 2004, the OASIS assignments would contain some previously unseen questions. For 2005 I also intended to provide different students with different assignment questions. The intention was that this change would reduce the incidence of students attempting to learn what questions were in the assignment by watching others doing the assignment, as well as reducing the effectiveness of the group approach to assignment completion.

The somewhat larger and weaker student group, the fractional reduction in course content, and the provision of different assignment questions to different students constituted the significant changes for this cycle of the action research. As in 2004, the primary research emphasis for this research cycle would be on collecting data to confirm earlier findings rather than on implementing change.

9.3.2 The OASIS assignments

As in previous years, only my part of the course was well covered with OASIS questions and could be assessed with OASIS assignments, of which there were two. However, as previously, my lectures did not start till week four of the semester. Therefore the OASIS assignments could not provide students with early performance feedback. The material of the first three weeks was assessed by a conventional test in week four, but the marked test scripts and results were not returned to students until the second half of the semester. The first OASIS assignment was held on the Friday of week six of the semester, the last day before the mid-semester break. Consequently, students completed half the course before receiving a significant level of objective information about their achievement in the course.

Just prior to the first OASIS assignment for 'Electrical Engineering Systems', the OASIS server had experienced serious loading problems during an assignment for a year-one Physics class of 470 students. These problems had led to the OASIS programmer modifying the software. Unfortunately, the modifications impacted adversely on the 'Electrical Engineering Systems' assignment. Some questions in the assignment did not appear for some students. Unfortunately, while some students had experienced similar problems during OASIS practice, they had not told us. Only two students informed us of any difficulties viewing questions, and they did so just the afternoon before the assignment. With 600 students busily practising on OASIS, just two such reports did not seem to indicate a significant issue with OASIS itself.

On the morning of the actual assignment, so many emails were received about the non-appearance of questions five and six that I reduced the assignment to the first four questions. The text of the email I sent to all students at 10:30 am, two and a half hours after the start of the assignment, is reproduced below:

> We have had some software problems with Q5 and Q6 so we have removed them from the assignment. The assignment now just contains four questions and your marks will be based on how you do or did on Q1 to Q4. Sorry about the strife that this has caused for some of you this morning.

Unfortunately, similar software problems also affected questions one to four, though to a lesser degree, and about 80 students could not view one or more of these questions. After the assignment had closed, I spent most of one day analysing the results and debating whether I could reliably estimate marks for questions that students could not see based on their performance on the questions that they had answered. I eventually judged that this would be very problematical. The situation was also made worse by the fact that the OASIS server had, without our knowledge or consent, been rebooted over the weekend and it was no longer possible to distinguish between a student who could not see a question and a student who could see a question but had chosen not to attempt it. There was also some evidence that a few students had been able to use their browser's back-button to make a second submission that had been accepted by the software. The decision was made to provide students with another opportunity to sit the assignment, and give them the better of their marks for each question on a question-by-question basis. This was explained to the students in lectures and the following email was sent to students:

> Some of you were unable to see all the questions that were in the first OASIS assignment. There's no perfect solution to this most unfortunate occurrence, but here is what we have decided is best to do about it.
>
> There will be a re-sit opportunity Monday 19 September, 8 am to 8 pm. The re-sit assignment will be the same as the first assignment, with all 6 questions, covering up to page 24 in the yellow course book - up to Q24 on OASIS. Please remember that, for each question in the assignment, there is more than one version.
>
> For your final result, I'll take the better of your two scores for each of the first four questions, and then add on your scores for questions 5 and 6 this time round. For example, if you scored 1, 3, 0, 2 on August 26, and then you score 1, 1, 2, 0, 2, 2 on September 19, I'll record your result as 1, 3, 2, 2, 2, 2 and you will have scored full marks.

Fortunately, the re-sit assignment ran very smoothly. However, considerable time was spent combining the results from the two assignments and conveying marks for all the questions from both assignments to the students so that they would feel confident they had been treated as fairly as the circumstances could permit.

The second assignment also ran smoothly, perhaps going some way to restoring the credibility lost by the first assignment's failure to perform properly. A few minor problems were encountered when students submitted answers such as 'a' instead of 'A', '12 V' instead of '12' (where the unit for the answer was already stated), and so on. Such discrepancies were resolved manually.

The assignments used both previously unseen questions and alternative questions. The number of alternatives for each assignment question ranged from two to five. While this number was not large, I wanted to be sure that the alternatives really were comparable in terms of content area and difficulty. I also judged that the number of alternatives would be large enough to make it difficult for students to familiarise themselves with all the assignment questions during the twelve-hour assignment period. While the introduction of alternative questions for 2005 might have made the assignments less susceptible to various forms of 'cheating', it made little difference to overall results. The average marks over a three year period for the two OASIS assignments in the year-one course are shown in Table 9.3.2.

Table 9.3.2

OASIS Assignment average marks in 'Electrical Engineering Systems'

Assignment	2003	2004	2005
First assignment	82%	75%	80%
Second assignment	92%	88%	88%

The somewhat lower average in the first assignment in 2004 was likely to have been a result of the slow speed of the system at that time, as noted in Subsection 8.3.2. Apart from that one anomaly, the averages showed a remarkable consistency over the three years for which OASIS assignments had been used in the course. The use of previously-unseen questions in 2004 and the use of both previously-unseen questions and alternative questions in 2005 did not adversely affect average marks.

9.3.3 Student voices: interviews and emails

The day after the second OASIS assignment, in order to obtain student feedback about OASIS, I sent the following email to students:

> I am keen to hear how you found using OASIS this semester. If you would like to take part in a short interview about OASIS, that would be excellent. Just reply to this email and we will organise a time. Preferably before the final exam so you can also ask me a few questions if you want. If you don't wish to be interviewed but could email me some of your thoughts and suggestions that would be much appreciated too. I am keen to hear your thoughts about OASIS.

The response to this email was gratifying, with over a dozen students providing useful comments about OASIS within twelve hours of the email being sent. Several students responded similarly the next day. In addition to these students, seven students were individually interviewed. From all this I gained a good insight into the way students went about their learning and, in particular, prepared for and sat the OASIS assignments. In this subsection, all quotes from emails are identified as such. The remainder of the quotes are from interviews and are not explicitly identified.

A few students commented on the software problems with the first assignment. These comments were supportive and suggested that we had handled the problem appropriately:

> I also thought that for a test that had been like a mis-test, the Department dealt with it in the right way: you couldn't lose the marks for those questions… if you got those [questions] right then you can obviously do them so you don't need to do them again. I thought that was a good way of looking at it.
>
> The errors with the first one [assignment] were slightly frustrating and, anyway, obviously beyond most peoples' control, but I think at the end it was well handled. I didn't really see any issue with having to re-sit it again. I only had to sit the last two [questions] again, so it didn't take much time. I thought it was a good way to handle it and I thought it worked well.

Much of the feedback paralleled that received via previous interviews and emails: it was confirming rather than groundbreaking. As previously, there were several requests for OASIS questions to cover more topics and perhaps partly replace the tutorials (each student had a one-hour tutorial each week):

> I found OASIS a good source of practice questions in addition to tutorials (which were quite limited). Perhaps practice questions for more topics could be placed on OASIS in the future. Thank you for your lectures this semester.
>
> [student email]

> I found it to be very useful as a practice tool. It would have been even better to have [the other lecturer's] material on there as well… I learned much from it. I personally think it could replace some of the tutorial questions (although not all) and hope all goes well with it in the future. [student email]

Students again readily voiced an appreciation of the instant feedback and the way questions could be repeated with different numbers:

> I think OASIS would be good for study actually, for exams and so on because you can just do lots and lots of questions, there's no limit to it.
>
> It was a very effective way of studying, with instant answers and many questions to choose from. [student email]
>
> It has abundant questions and it tells the answer instantly. I think this is the neat thing about Oasis. [student email]
>
> It is an excellent tool for practising questions as it gives much more variation in questions and the correct answer is received much more quickly compared to the more 'traditional' way of practising problems using the textbook. [student email]
>
> You could practise it and practise it so you know how to do it. It's an unlimited supply of questions which is quite nice. It doesn't help you with the concept stuff but it helps you with the question stuff so you make sure you know how to do it.
>
> Obviously the limitation of the textbook and the tutorial ones [problems] is that once you've done all the ones, say, about transformers, there aren't any more, whereas with OASIS, to repeat the same process of calculating over and over again, it cements it in your mind.

As the last comment suggests, repetition of questions was seen as way to learn and retain skills:

> I think one of [the questions I did] up to about 10 times…I think I probably learnt more from the questions that I got wrong the first time because now I know how to do it whereas the ones that I was a bit shaky on I still got it right. I probably might not be as confident in that in the future.
>
> No, I didn't write any notes. I thought if I do these enough times I make sure I know how to do it. It will be easy. I just wrote workings.

Some students did spend considerable time preparing detailed notes on the OASIS questions so that they could refer to them later:

> S: If I still can't find [how to do a particular question], then I'd probably just flag it and move on to the next one and then ask someone else on the day if they could do that one and then if they could I'd use their book instead of my book to do that question.
>
> CS: So it sounds like quite a few of your friends are actually writing solutions down in books?
>
> S: We all write down in a book to show how the thing came out. We write the question down first so it showed it all nicely. It just made it quicker and easier.

Generally, students displayed good judgment when it came to assessing their own learning and grasp of the material. Their judgment in this regard determined timing as well as the amount of repetition and note-making:

> I did leave it fairly late… I figured there was no point in doing it weeks out because the content was the same as we had already done for the lectures so it wasn't like I had to practise to learn stuff. The practice for me was just to get myself accustomed to what the questions would be, and how they would look, so when I got in and only had an hour, I wouldn't sit there thinking 'this is something completely new', I'd sit there thinking 'yeah, I've already seen this one before, that's fine'. It wouldn't necessarily be that I went and practised and practised until I knew how to do every single question off by heart or anything like that. I just practised enough so that I knew what they would look like, where the questions would come, the sort of answers that you had to give because the concept of how to do the questions came through the lectures and the tutorials.

While some students aimed to start work on the questions well before assignments and to refresh themselves with a second working of the questions immediately before the assignment, the effects of time-pressure (from the demands of other courses) were clear, particularly in the case of the second assignment:

> I try to do them about a week before… I just do all of them all over again the day before the assessment. Just to remember.
>
> For the first one I spent a few hours and did half the questions about 10 days before and then a few hours a week to do the rest of them. Then a couple of days before I just worked through them each again to make sure I hadn't forgotten anything… [It was similar], but with less time, for the second one.
>
> CS: How long before the assignments would you have started looking at the OASIS questions?
>
> S: For the first test [assignment] probably two weeks… I finished them all with a couple of days to go… but for the second test I did them all in one day because I was busy doing a design project.
>
> Doing a conjoint you're always really busy… So in the last three or four days I sat down and had a look and made sure I could do it.

Clearly, the assignments were good motivators of student practice. Some students mentioned this explicitly, with one student even stating that she would prefer tests over assignments as the former were more motivating:

> I thought Oasis was very helpful as it was a source of renewable questions with a different answer every time. The assignments also were a good motivator to study the material. [student email]
>
> Maybe it would be better if it's a test for me because I think tests help us more, because we know under the stress and all the pressure we will learn more, that's what I feel.

There were, as usual, several requests for hints or model answers, particularly when students were not near the correct answer. Again, this was tempered by the understanding that such a system could be abused:

> Sometimes you do them and you find it quite difficult to find where exactly you went wrong, especially if you're a long way off. If you're close you can sometimes just find it.
>
> Maybe if some of the harder parts of questions could have a page of explanations for the answers, as if I couldn't work things out I tended to go through a trial and error to work out how to get to the answer -not very scientific. [student email]
>
> After the third or fourth attempt that someone's obviously not getting it, it should come up with a hint… [However] if you have it [a hint] coming up after the third or fourth attempt, I guess the rumour will spread, people will just click submit, submit, submit three times…
>
> From a learning point of view, I think having no help, or having little help, is the best solution because it makes you, if you can't get something right, go to the text book, lecture notes and find how to do it… but from an assessment point of view you just want to know how to get the right answer.

Overwhelmingly, the feedback received about OASIS as a practice and learning tool was positive:

> I don't really have a fault with OASIS; I think it's really good software. For what I'm after it delivers everything.
>
> I thought oasis was an excellent way to learn the concepts taught in your lectures. Not only was it a good way to sit our tests but it was a valuable learning experience as well. [student email]
>
> Last semester when I was using it for Physics it was really good for studying with because there was one thing, thermodynamics, that I couldn't understand at all. I went to OASIS and kept on doing the questions and then I'd go to the textbook and look one thing up, okay it works. That would help me to actually understand the material and by then you actually know it quite well.
>
> First of all I think OASIS is an excellent online resource because all the other online things that come with these text books and things don't work… [OASIS] has the advantage over the other programmes that you get with these textbooks: it gets to the point really quickly. You don't have to read a story…
>
> A very effective learning system for practice questions… the system itself is a new form of education that has potential to be very effective in years to come. [student email]
>
> I thought Oasis was great! Brilliant way of learning - being able to go back and repeat it with different numbers to make sure you understand. I thought it was a really good system! Great way of revising. Don't know that it's such a good way of assessing only because people just get help from others! Unless they were supervised assessments. But yeah, I really enjoyed Oasis! [student email]

The last email typified much of the feedback: students appreciated OASIS as a practice and learning tool but had reservations about its use for assignments. Much of the feedback favoured supervised assessments rather than unsupervised assignments: the former were seen as being fairer. Part of the problem appeared to be that the two assignments contributed 10% of the final course mark. Some students felt this weighting was too high.

> [OASIS] was a very effective learning system for practice questions, but not so much for assessment. Some people who did not know what they were doing would ask people that knew, and then they did it for them. It was not fair to some students who sat the assessment with no help from anyone else. If stricter guidelines were set, with supervision being a possible implementation, the OASIS assessment would work well. [student email]

> I've noticed that people who don't bother studying for it get excessive help from their friends and get higher or similar results to those students who studied hard for it. This is a bit not fair. What I would suggest is that OASIS could be done under tutors' supervision… [This] would force those students who rely on their friends to study and prepare for it properly and so I reckon these students will even do well in the final exam. [student email]

> You could get a student who's really struggling with the stuff but they have ethics in that they don't want to cheat so they'll do it by themselves whereas there may be other people struggling who just get a smart friend to come along who gets the 10% the other person doesn't and that's an unfair advantage so I'd say it's better to probably book the computer labs and have them supervised.

> Although I found it a very useful learning tool, as an assessment tool it was poor. It did not seem fair to offer a total of 10% for a test which students who knew nothing could easily ask for help on while those who gave it an honest effort risked their marks. [student email]

> If marks are going to contribute to final year marks, it should be done under TEST conditions; this gives each student a fair chance of showing how well they know the course, not his friends that were beside him helping him with each OASIS question during assessment (for example)… assessments should either be done in fair test conditions because it contributes to final year marks, or still continue to use OASIS for assessment but not contribute so much to the final year marks… [student email]

It was pleasing that no students referred to spreadsheets. 'Group assignments' were certainly still in evidence, in spite of the alternate questions in the assignments which rendered such practices less rewarding.

> One of the things I didn't like about OASIS was the fact that it perhaps wasn't fair from the point of view of those of us who put in the effort and tried to do it and worked hard to get all the marks compared to those people who perhaps worked with three other people on their assignment.

> People cheat in the assignments… You see about four people in a computer lab and they go through and do each one's assignment.
>
> During OASIS time they just work as a group, everybody do one person's assignment, everybody calculate at the same time and they just put the answers which have the same answers so everyone got good marks apart from me. I don't know if I'm stupid or I just want to do it by myself but sometimes I just didn't get good marks… it's not fair.

Based on the student feedback I received, I judged that these formal groups had become less popular and that to some extent they had been replaced by less formal arrangements where students might ask those around them for help as they needed it:

> You go into the computer labs and everyone just sits round the computer and does it together… if you don't know a question it's so easy to go to the person beside me and ask: 'how do I do this?'
>
> I don't think you can stop people helping each other… It's human nature to help people if you can do it…. It's a really difficult one trying to get people to be honest and do it by themselves but you're never going to.
>
> [Other students] came and asked me for some help and I just instinctively went and helped them.

It was clear that a significant number of students considered themselves to be disadvantaged by the collaboration of others in assignments. However, some students who had collaborated stated that no great advantage had come of it:

> I did my assignment by myself but… other people… asked me for some help… I didn't really see it as cheating. I guess the problem with any closed book exam is that it's not a very real world kind of situation. Its not like now you're an engineer, off you go, like you have to do this and you certainly can't look at any text books or go on the Internet or anything like that, so the fact that you can ask other people is probably not an issue because it's still easy enough for either the other people not to know or to make a silly mistake that you should have just checked yourself.
>
> I think there were about 12 of us… we had only six people on the computers at once, we had people behind us checking our answers and making sure that we were getting it all right. I ended up getting it all wrong through my own fault, but there were always people making sure that we weren't [making mistakes]… We all do the study for it on our own and I did all the practice questions and I did them all in my book so when I came to do the test I looked for the question that was similar and I could run with that because I already knew how to do them, it was just people behind us were checking but the study for it was really good and I think we would have got pretty high. If you do the study for it I'm not sure you really need the people behind you.

While the students were, in general, enthusiastic about the way OASIS had helped them learn and understand the material, there were occasional suggestions that some

learning was surface rather than deep learning. One student actually voiced concerns about the depth of her own learning in an interview:

> Sometimes you just memorise all the calculation, all the steps, like for me, I'll forget why I get an answer but I just know how to do it, to answer the question, I know how to do it because I know exactly the calculations to use. What formulae to use. Afterwards I forget… One thing I like about OASIS is that the practice questions are really similar to the actual assignment we are doing so we know exactly what we are going to get, but that's pretty much what we did in high school because the high-school teachers told us what to do and we know what we're expected to do, but I think for university maybe we have to, there's more. I don't know if it's good or bad for us. Maybe it's there for our points and to get our degree but I like to try things that I don't know. I suppose I want to do things that I don't know, like I want to do things that count towards my own knowledge. If I'm told what to do then I do the same for tests and I finish university but I don't know anything… I actually struggle with some of my papers but I want to, I like to, learn.

Overall, the Year-one feedback was positive. It confirmed that OASIS had supported student learning well. However, a significant number of students still felt that the use of OASIS as an assessment tool was inappropriate, even though there were now alternate assignment questions as well as individualised versions of each question. I judged that this issue would need further addressing for 2006.

9.3.4 The final examination

The final examinations in 2003 and 2004 had consisted of six questions. In 2005 it was reduced to five, a result of the removal of one part of the course. Apart from this reduction, the examination followed the previous pattern, with two questions being contributed by myself. These two were supported by OASIS questions while the other three were not. The average marks for my two examination questions were 68% and 52%, respectively, as set out in Table 9.3.4. The average marks for assignments, tests and examinations over the three-year period for which OASIS assignments were used in the course are also shown in Table 9.3.4.

I was pleased that the 2004 improvement in the result for the first examination question had persisted in 2005. This improvement was discussed in Section 8.3.6. However, I was most disappointed by the significant decline indicated by my second question, from 65% to 52%. I judged that this was a result of several factors working together. First, I had been required at short notice to take an overseas trip that necessitated reducing the set of lectures related to this question from eight to seven.

Table 9.3.4

Assessment average marks in 'Electrical Engineering Systems'

Assessment	2003	2004	2005
Assignments (two)	87%	81%	84%
Tests (two)	62%	58%	55%
OASIS-supported examination question one	54%	71%	68%
OASIS-supported examination question two	68%	65%	52%
Examination	58%	60%	58%

Second, this set of lectures had come later in the course in 2005, a time when the students were preoccupied with assessments in other courses. Third, although I had tried to maintain a consistent standard, it may have been that the question was somewhat harder than in previous years. Fourth, there had been a significant decline in the ability and achievement of the group of students taking the course in 2005. Reasons for this decline were advanced in Subsection 9.3.1. This decline did not reveal itself in the tests and assignments which were based on relatively small sections of the course. It also did not reveal itself in the examination average mark, perhaps because the question that was deleted from the 2005 examination had returned a below-average result in 2004 and 2003. However, the decline did show itself in the final course grades: in 2005, 74 students received a D or a D− grade. In 2004, only 40 students had gained such grades. It almost was as though the increased student numbers in 2005 had simply translated to a corresponding increase in the number of D and D− grades. Similar results were observed in other year-one Engineering courses. Concern in the Faculty over the relatively poor performance of the 2005 year-one group led to significant changes in entry procedures for 2006. In 2003 and 2004 the Faculty had accepted approximately 550 students into its first year. In 2005 this number had been increased to 600. In 2006 it was reduced dramatically to 450, with a further 200 students promised a chance of Semester-two entry should their results in Semester one in certain prescribed BSc courses be deemed adequate. Of these 200 students, only 30 subsequently met the standard.

9.4 OASIS usage in the Department of Physics

During the first semester of 2005 the Department of Physics used OASIS with two classes of around 250 students each. From the third week of the semester, they had weekly OASIS assignments. The data from these showed that, surprisingly, the performance of the Physics students differed significantly from those studying Electrical Engineering. There were relatively large numbers of students who did not submit any given assignment: in one course of 280 students, 40 students did not submit the first assignment while 60 did not submit the last. There were also quite a few students who answered only one assignment question. Failure rates approached 50%. These statistics were most surprising: the assignments typically comprised four questions chosen from six, all of which had been previously available for practice during the week before the assignment. By contrast, in Electrical Engineering, the failure rate was nearer 15% and very few students did not submit any given assignment, perhaps six or seven in a class of comparable size. The Physics students also had ample opportunity to complete assignments: the Physics assignments were available through a 24-hour period, twice as long as the period used in our own assignments. Nor were the OASIS assignments undervalued: a typical Physics course had four OASIS assignments that contributed a total of 15% to the course grade. By comparison, the ECE year-one course 'Electrical Engineering Systems" had two assignments worth a total of 10%, while the year-two course 'Circuits & Systems' had four assignments worth a total of 10%. The fact that the assignments were based on only six questions, all previously available, provided students with a good opportunity to prepare in advance, perhaps through writing out the method or constructing a spreadsheet. Certainly, there were a number of students who scored full marks after 10 or fewer minutes. However, these students were a small fraction of the total. Much greater numbers of students gained scores that suggested they had hardly prepared at all and consequently grasped the material poorly.

I was certainly surprised and disappointed by the results gained by many of the Physics students. However, the instructors for the year-one Physics courses did not share my disappointment. In fact, they considered that OASIS was well-used and well-liked, and that it had lifted student achievement levels significantly: "Overall positive student feedback and a noticeable improvement in student problem-solving

skills in the first full year of OASIS assignments are very encouraging" (Kruhlak, Mullins, Coghill, & Ng, 2005b). Student surveys conducted in the Physics classes certainly did show that students appreciated OASIS. For example, in a survey of one class, 80% of the Physics students agreed or agreed strongly with the statement "The OASIS system (on-line practice followed by assessment) gives a better method of coursework assessment than take-home assignments" while only 6% disagreed or disagreed strongly. For the statement "OASIS measured my learning in this course fairly" the corresponding figures were 68% and 7%, respectively, while the statement "OASIS helped me improve my problem solving skills" returned figures of 70% and 8%. Only 3% of the students surveyed judged that the instructions for using and accessing OASIS were not satisfactory; and only 2% stated that the allocated period of one hour was not long enough for the assignments. The surveys also included three free-response items as follows:

> What did you LIKE most about the OASIS system?
> What did you DISLIKE most about the OASIS system?
> Please write any other comments you wish to make below.

In response to these items, students commented that they liked the instant feedback and the ability to practise the questions. They did not like the lack of credit for working, or the lack of hints and model answers, or the fact that some students could and did cheat. Students also stated that they wanted more OASIS questions.

The results from the Physics surveys were remarkably similar to the results from surveys of our own students. For example, 71% of students agreed or agreed strongly and 9% of students disagreed or disagreed strongly with the ECE year-one survey item "OASIS helped improve my skill level". By comparison, the Physics year-one survey item "OASIS helped me improve my problem-solving skills" returned figures of 70% and 8%, respectively. Further, the positive comments about instant feedback, the negative comments about the lack of hints and model answers, and the requests for more OASIS questions, mirrored exactly the feedback we had received.

Data collected automatically by OASIS also revealed similar usage patterns to those noted in ECE. Students did use OASIS extensively for practice. One week before each assignment, six practice questions would be released. On average, each student would practice each question four times or more, with a student-participation rate of

over 90%. It was also clear that, on practice questions, students tended to score either full marks or zero marks. Any intermediate marks were much less common, even for questions scored out of 10 with half-mark increments. Physics instructors judged that many of the zero scores were produced when students wished to print a copy of a question and submitted a null answer so as to gain the answers to the question they were printing. A typical spread of marks is shown in Figure 9.4.1.

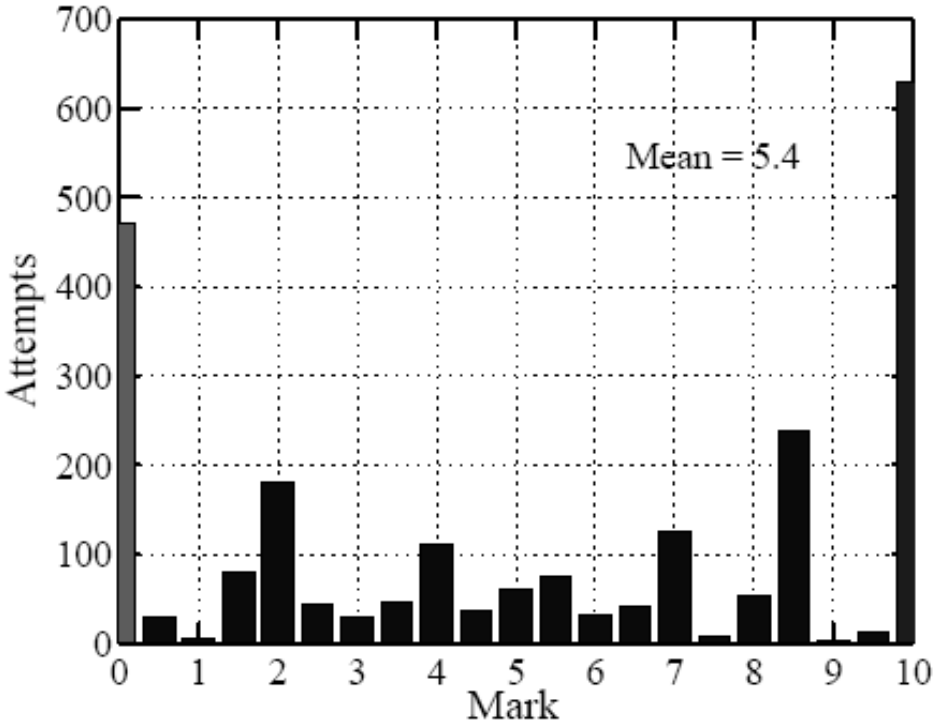


Figure 9.4.1: The mark distribution produced during practice on a typical OASIS question by students enrolled in the year-one course: "Physics for the life sciences". Adapted from Kruhlak, Mullins, Coghill, & Ng (2005b).

Again, Physics students tended to leave their practice till the day before assignments, a trend that noted in ECE. The effect in Physics was perhaps more pronounced for two reasons. First, the practice questions were released only one week before the assignment. Second, the number of questions was small, so students could perhaps expect to master them in a day. For my assignments in ECE there were far more questions and the assignments typically used previously-unseen questions, so more preparation was appropriate. Student activity in a year-one class through a semester is shown in Figure 9.4.2. As expected, the graph shows major peaks prior to assignments. The peaks are higher at the start of the semester, and at the start of the

second half of the semester. This is to be expected; at these times the students would have been relatively fresh and they would have had fewer assessment activities in other courses. The graph also shows smaller peaks in student activity prior to tests and the final examination.

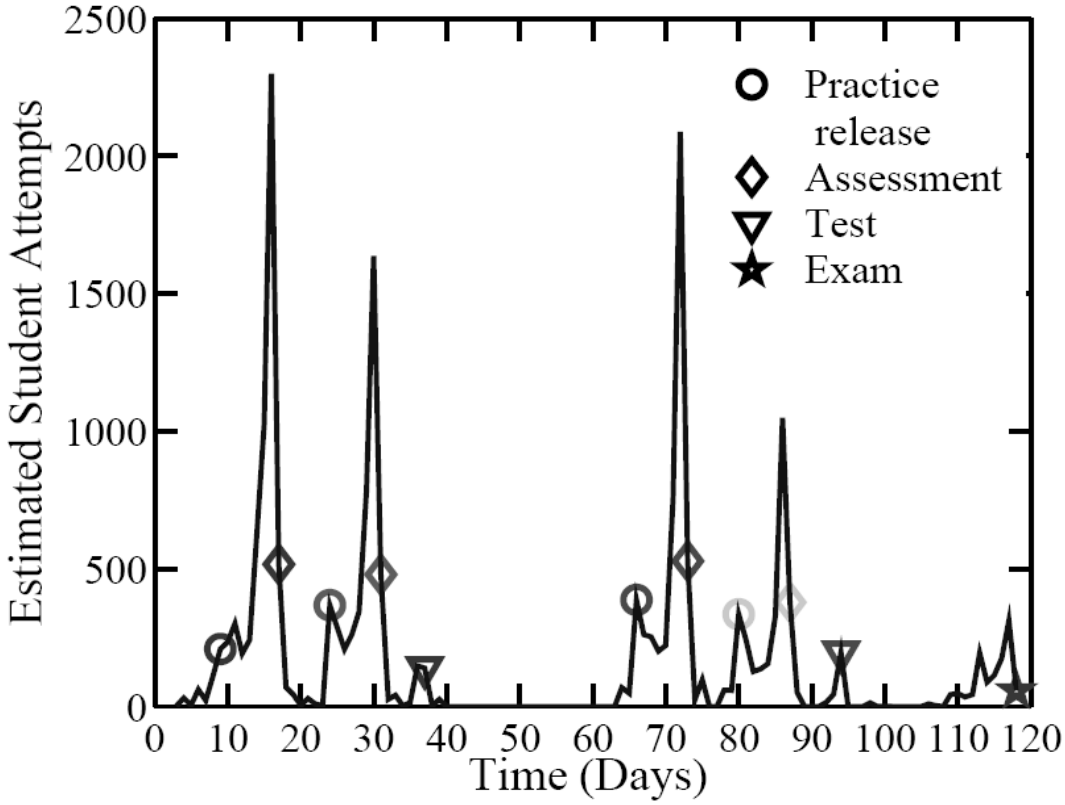


Figure 9.4.2: Student usage of OASIS during one semester by students enrolled in the year-one course: "Physics for the life sciences". Adapted from Kruhlak, Mullins, Coghill, & Ng (2005b).

The instructors involved with the Physics courses were certainly enthusiastic about OASIS: since the 2003 trial, its use had grown dramatically with more than 1000 Physics students using OASIS in each of 2004 and 2005. Since, for each of these students, four online assignments had replaced two written assignments, the implementation of OASIS had eliminated the marking of literally thousands of assignments. Instructors were happy about this, stating they preferred creating OASIS questions to marking scripts. Students were also receiving more frequent and prompt feedback.

My concerns about the relatively low scores on the assignments were explainable, at least in part, by the reality that the Physics courses had much lower entry requirements than the Engineering courses with which I was comparing them. However, some Physics instructors had also voiced concerns about the relatively low marks in the Physics assignments, particularly those gained by the Physics students hoping to continue in Physics. In comparison with those students, the Physics students hoping to qualify for entry to the Medical School appeared to be more motivated, submitted around twice as many answers to OASIS questions during practice, and scored more highly in the assignments. This latter group was described as producing "impressive outcomes" in their assignments, while the former group of students, those hoping to major in Physics, was described as producing "poorer than expected outcomes"(Kruhlak, Mullins, Coghill, & Ng, 2005a).

In summary, the Physics experience was, in general, similar to the ECE experience: both instructors and students considered that OASIS had lifted the students' problem-solving skills, students preferred OASIS over more traditional alternatives, and instructors judged that their workload had been reduced.

9.5 Feedback from postgraduate students

So far my interviews with students had focused on experiences they had had with OASIS only shortly before the interviews. I was also interested to gain insight into how students perceived their OASIS experiences a few years later. Perhaps with the passage of time their perceptions would change. In particular, perhaps they would find that the knowledge or skills they had gained through OASIS practice was not long-lasting. In order to investigate the long-term implications of OASIS usage, I emailed all current ECE postgraduates. Subsequently, I received 10 emailed responses, some of which were quite lengthy and were clearly the product of considerable reflection. I also conducted one interview. In this subsection, unless otherwise noted, the quotes are from postgraduate emails.

The feedback generally aligned well with undergraduate feedback. In some cases there were mismatches because the postgraduates referred to circumstances that no longer pertained to the current OASIS implementation, or to the way in which I was

using OASIS myself. For example, some students voiced displeasure at being required to use a software calculator for OASIS tests:

> It was a disaster when we had to do tests in OASIS [because] students were required to use the Microsoft windows calculator (that you find under accessories). This wasted a lot of time, for instance finding sine, cosine functions etc. I believe students can now use their own calculator.
>
> Initially, the programming of solution equations into calculators resulted in students only being able to use the Windows calculator which was clumsy and irritating... Given that these errors were more likely to occur with the Windows calculator, this sort of levelled the playing field between more and less able students. I don't view that as a good thing from an assessment.

The postgraduates, as they reviewed their undergraduate years, were ambivalent about the use of OASIS for tests. They had not experienced OASIS assignments. However, they were positive about the use of OASIS for practice, particularly appreciating the instantaneous feedback and the question repeatability:

> In general OASIS is an excellent tool for practising; perhaps tutorials should all be based on OASIS.
>
> OASIS was useful for practice due to the instantaneous feedback as to whether you were right or wrong. In general, if my memory serves me, it also provided a wider range of questions than would usually be provided on a tutorial sheet. OASIS also provided a good opportunity for students to work together as groups on problems they would individually find too challenging.
>
> It generates random numbers for the same questions, so that that when the students try to work out the solution, they can try the same questions for a few times and ensure that they didn't get the answer correct just by luck. I also like that the questions on OASIS are set from easy to hard and usually the theory/methodology from the previous questions can be used in the questions later on.

The postgraduates did, in general, see OASIS as good for learning. The idea that the ability to repeat questions with different numerical values enhanced learning came through in a number of emails.

> I found being able to repeat the same problem with different numerical values valuable because it meant I could repeat that problem several times until I was confident that I fully understood the underlying concepts and had developed the skills to solve that particular problem structure. i.e. re-solving the same problem as if you'd "never seen it before" in terms of not knowing the numerical answers means the student can really learn that particular problem-solving strategy and make sure they properly understand the related concepts.
>
> CS: In terms of learning, for what did you find OASIS particularly good /bad?
>
> S: GOOD: circuit-solving type problems, where repetition of the same problem structure is particularly useful and is perhaps the only way to

> gain a confident grasp on the material (in the same way as Maths can only be learnt with practising problems, never solely by reading mathematics textbooks). [nothing in BAD category stated].

However, the students were also aware that some students adopted a surface-learning approach, simply memorising a formulaic method for each question:

> I think some students might tend to memorise the equations needed for each question and simply dump them out during the test. Therefore I believe only a minor portion of marks should be allocated to OASIS test questions. (Say 15% of final, for example)
>
> I know that many students just memorized the questions and the ways of solving them, so they didn't really understand the question or the things that the lecturer wanted us to know.

There was some recognition that OASIS could support both surface-learning and deep-learning approaches. It was a choice that each student could make:

> Some people just blindly memorise the method but also the good thing about it is you can just print questions out, you can do some working on the side and you actually understand what is going on… So I think it is both good and bad.
> [postgraduate interview]
>
> I didn't really like OASIS during my undergraduate degree. Since, as far as I can remember, people tended to memorise steps to solve a particular [question] whether those steps were making sense or not, and I was one of them… But on the other hand, working out the steps to the solutions helped me a lot to think/derive-in-the-reverse the things going on in the question, and more exercises of OASIS gave me a chance to improve my problem-solving skills in the long run.
>
> At the end of the day, OASIS forces people to either adopt a short-term strategy of solving specific problems or a long-term strategy of solving any problem. The less able students will tend to fall into the former group, the better into the latter. As a result, better students may well develop skills providing a solid method for tackling problems using theory. However, retention of specific topics probably depends on the individual, at what stage in the degree the material was covered and how much it was reused throughout the degree. Poorer students will still just cram for exams, especially as long as they are paid off for doing so with easy marks.

In my email to the postgraduates, I asked a number of questions in order to focus their responses. In particular, I asked the following question at the end of my email:

> When you learned skills, etc. with the help of OASIS, did those skills, etc. stay with you or did you lose them soon after? (This question is motivated by the fact that sometimes we cram for an exam but what we cram is gone a week later).

Generally the response to this was favourable. However, students were aware that the answer depended on a number of variables, including: what depth of learning the students had chosen, how much practice they did during the learning process, and how often they had reinforced the skills since first learning them:

> No, I didn't lose them after I had done the test or practices, but I know many of the students did… many students just memorize the questions and the ways of solving them.
>
> I think that the skills learnt with the help of OASIS, and the resulting elucidation of related concepts stays with the student. The specifics of how to solve a particular problem would naturally be forgotten quicker, but even this would stay with me for at least the duration of that semester.
>
> How much I learnt and kept really depended on the circumstances surrounding the assessment (was I busy with other assignments etc. at the time), how many problems were provided to cover the same general theory, how interested I was in the material and how much it was used throughout the course of the degree. For example, op-amps, BJT biasing, mesh current & node-voltage analysis, and general circuit analysis techniques stuck with me quite well, such that despite the fact that I haven't used a BJT as an amp since second year I still remember how to do it with a little memory jog. However, I have almost no recollection of the power systems stuff that was assessed through OASIS. I have a vague recollection of some control systems stuff, but that is mostly due to similarities of repeated material throughout the degree.

In summary, the feedback from postgraduates generally paralleled that from undergraduates. No conflicting evidence was produced. Again, OASIS was seen as a valuable tool for practising skills, and the instant feedback and the repeatability of questions were appreciated. The postgraduates saw OASIS as less useful for tests than practice, but some of their reasons were not applicable to the newer implementation of OASIS. For example, the use of the Windows calculator did not apply, nor did the way in which students could simply memorise the solution methods for questions (because assessments now often used questions that students had not previously seen). However, their observations that some students prepared for OASIS assessments in a shallow way while other students made a genuine attempt to master the material certainly were still relevant. This theme had also emerged in feedback from undergraduates, the difference being that the postgraduates were more aware of the choice confronting students and were able to articulate it explicitly. It was pleasing that the students judged they had retained quite well the skills learned through the use of OASIS. This judgement was consistent with staff observations on the same matter.

9.6 Interviews with academic staff

Although I frequently had informal chats with staff about OASIS, there had been only one round of formal interviews with academics involved with OASIS, and that had taken place in 2003 (see Section 7.8). These interviews had been conducted by a member of CPD. Since these interviews, OASIS and its implementation had changed significantly and the relevant staff group had also changed considerably: some staff who had used OASIS were no longer employed by the University, while other staff members who had not used it in 2003 had since adopted it (No staff members who had previously used OASIS and who still remained on the teaching staff had ceased to use OASIS). Further, since 2003, staff had also had time to develop a longer-term perspective. For these reasons it was deemed important to have another round of interviews. To this end I interviewed several staff members. To obtain the maximum benefit from this exercise, I focussed on staff who used OASIS in courses other than those I taught. Some of these staff had used OASIS since its inception, while others were recent adopters. All interviews were recorded and transcribed. Unless indicated otherwise, all indented text in this section consists of quotes from the interviewed instructors. Instructors may also be denoted by I. The interviewer is denoted by CS. Further details concerning staff interviews were given in Subsection 5.7.6.

There was a considerable variety in the reasons that instructors gave for adopting OASIS in their courses. One staff member even reported that he had been coerced by colleagues with whom he shared a course to implement OASIS. Fortunately, he did also report that he was glad he had been pressured to use it:

> Why did I use it? Well I was pressured into trying it and I'm quite happy that I was pressured into trying it because I've actually found it extremely worthwhile in that section of work and I'm more than happy to actually try it now in other sections of work… I'm not unhappy with the fact that I was sort of pressured into doing it; I was impressed with its capabilities.

A number of other instructors stated that they had adopted OASIS in order to address the problem posed by large numbers of students in their classes who did not put adequate time into skills practice and problem-solving. It was also noted that this problem that had become worse with time as class sizes increased and the study habits of the student population changed:

> Over the years, the numbers of students kept on increasing and the quality… continued to decline… more students at a lower level failing more and it just wasn't succeeding at all
>
> Some years ago I was concerned that students didn't seem to have the skill level that they should have… though they thought they did… we'd always believed that students would start to understand things better by working problems but they never seemed to be keen to work problems… The number of students that did all the problems in a text book was vanishingly small and the odd one that did… never got it right. So OASIS was founded on the basis that maybe we could do something about these things
>
> If you're trying to meet the needs of the class you've got to try and tailor it to the average student and the average student changed… there is an element of 'if we're not given it in lectures we're not going to study it, we don't read the books, we don't do the problems in the back of the book, we don't purchase the book'… it's certainly notable from a class I might have had a decade ago to a class that I have now. There was a slow tailing off to the point where you really couldn't assume that students would do background reading or the questions in the back of the book… When we had smaller numbers… I would get the students doing tutorial problems on the computer and solving them… Then the class grew too large; I couldn't do that, so it became chalk and talk on the blackboard. At that point I found very quickly that the students weren't really doing the problems, they were waiting for the tutorial, they were all at the tutorial, they loved the tutorial but they really weren't looking at the problems… I originally had this computer programme that I thought they'd make use of but they didn't… then I moved across to MatLab and used one of the packages in Matlab… but even then they weren't really attacking it. I could show them in class but… they weren't really doing it… My main motivation for using OASIS is to get students actually working during the course and learning.
>
> I: So OASIS was there to improve their skills… it was used to ensure that students would practise some of the design problems, solve equations themselves and consequently learn EDT [Electronic Design & Technology] at the same time…
>
> CS: How do they normally go on the OASIS part of the test?
>
> I: Oh usually quite good. OASIS is always a saviour. If it wasn't for OASIS my part would get horrible statistics.

Unsurprisingly, OASIS was used in different ways in different courses. Since my arrival at the University in 2002, I had used OASIS assignments in the three courses I taught. These courses had between two and four OASIS assignments, depending on how much of the course was covered by OASIS. I considered four OASIS assignments appropriate for a course fully covered by OASIS. The Physics Department had adopted a similar four-assignment model but decided to implement weekly OASIS assignments in 2006 to motivate more consistent study habits amongst its students and to provide students and instructors with regular feedback:

> Physics have now switched from four tests per course spread over the semester to ten tests per course spread over the semester… to get more regular feedback from the students about how much understanding they're showing when doing the course and that's fantastic. That's what OASIS can do. The marking load completely disappears so use it more.

Other instructors, particularly at year-three level where class sizes were more manageable, had implemented supervised tests rather than unsupervised assignments. There were typically two or three such tests. These tests were often partly OASIS tests done on a computer and partly written tests. The half-computer, half-traditional approach was adopted for two reasons. First, it enabled the test to examine a broader range of skills. Second, it made possible OASIS tests for classes where the number of students exceeded the number of available computers: for the first half of the test, one half of the class sat the written part while the other half sat the OASIS part; in the second half of the test, these two groups would swap activities and rooms:

> I didn't want to run it [the test] purely in OASIS, I wanted to have a written section that could address… conceptual issues.
>
> A test at a lower level might be almost fully OASIS. I don't believe that's a good idea in later years because you want… more of a qualitative assessment rather than quantitative in terms of numbers... I want to get a feeling of their understanding.
>
> Because of the class size we found that we had to run half the class doing the written section and half the class in the computer room and then swap them over. Actually by and large everyone was happy with that. It worked pretty well. The students were in a hurry to get down there and log onto the network … so they don't have time to pause and stop and talk and it worked really well.

The OASIS part of the test would be automatically timed out. This was also required for half-and-half tests even when there were enough computers for all the students and no room change was necessary: otherwise some students would spend virtually all their time on the OASIS half of the test, believing that they could gain enough marks on that half alone to pass the test:

> Some of the students chose to spend the full 50 minutes only on OASIS and they were banking on getting 100% of that mark… To defeat that… I put a timer on the OASIS so… they don't have the full time and they have to spend time on the other… I don't know if that has eliminated it or not because when you examine the results in the written part, there's usually a wide spectrum right from poor to fantastic.

The instructors that had adopted OASIS for their course considered that the students, in general, liked it and engaged well with the material:

> The students do seem to like OASIS-type problems
>
> OASIS, I think, provided them with something like a dozen questions in that topic and a large part of the class did seem to actually do those problems. One could look at the statistics from OASIS… there was considerable engagement.
>
> I think it's been very successful. I'm not aware of anywhere where it would have hurt at all but the students really use it a lot and I think that's a sign that it's working.
>
> Lots of students use it, especially leading up to assessment but even just for general practice, not their assessment, and just talking to students they're quite happy with that, they enjoy using it.

Reasons given for the good level of student engagement with OASIS included the instant feedback and the repeatability of questions, as well as the fact that the questions could appear in an assessment:

> The students do seem to like OASIS problems in that they do get immediate feedback.
>
> I think student learning has improved if only because they seem to do the questions…I think they could always have done questions in the back of text books and in optional assignments but they just didn't. I think the marking feedback is the key behind that one. I think 'try and find out straight away what you got right or wrong' is the real key there.
>
> When they start doing them [OASIS questions] their ability to do these problems is actually very, very poor. They would have a success rate of about 20% or something like that, they'd get them wrong, and they, through their own efforts, bring that up to the high nineties… I've seen students who say 'I keep doing this problem until my average score is 80%'. I said 'why?' 'well, I just set that as my own personal goal that four times out of five if I did this problem I'd get it right' but because they got so many wrong at the beginning what that actually means is that at the end they're getting it right nine or nine and a half times out of ten and that's actually quite different. You could never do that with a text book. You could never do that with a test system and if you set the students that as a goal and said 'I want you to do this', as a prescription for the course, that would be draconian, but students will impose that restriction on themselves and they'll do it and it's really quite remarkable, they do get very proficient at doing these things.
>
> [OASIS] forms a really good place in terms of assessment because it gets the students trying problems and they don't try them if they don't think they're going to be examined. So I have to tell them they're going to be tested on these. Any time I tell them this won't be in the test you just notice the fall off in terms of trials.

The instructors reported that the implementation of OASIS had raised student skill levels and understanding:

> Skills require practice, OASIS provided an environment for them to practise and certainly the skill level went up. I think that the understanding level went up.

> I asked some fairly tough questions in that OASIS test and the vast majority of the students actually performed surprisingly well, much better than I had imagined, and when I examined similar material in the final exam I also observed a similar improvement in terms of understanding and familiarity with the material. So I actually found it pretty positive.

> I did observe what I viewed as a noticeable improvement in terms of the students' ability to answer that type of question in the final exam and I assumed, in fact I can be quite positive about it, the improvement was linked back to their time and their engagement on OASIS, not much doubt about that.

> The students certainly showed signs that they had a better understanding of the material… it's a subtle thing, if you're marking an exam and you think the students are doing a bit bad you're looking for marks, whereas if they've done really well you tend to be looking for parts they've not really answered well so you can take marks off. I found myself shifting from looking for marks to looking for 'really did they understand it well enough' and I would remove a few marks if they didn't show that understanding. To me that's an indication of a class that's getting to grips with it better.

> It's quite dramatic. It's completely turned around. They can do these things. They can do them and they're skilful at it.

The ways in which students approached their studies also appeared to change somewhat. There was now less emphasis on obtaining the right answer and more on knowing the right methods. This was perhaps to be expected, since, with OASIS, the right answer kept changing. Also, after using OASIS, the students were judged to be less dependent on the lecturer for guidance with their learning:

> They now come and ask… 'what methods should I use for problem seven because I can't consistently get the right answers and my friends can't either so we must have the wrong method'. This was such a great change… we could talk about four or five different methods and they'd go off and try all four or five methods and it was absolutely fantastic.

> I do observe that the students appear to be working amongst themselves… the class becomes its own teacher to an extent but obviously the OASIS environment has been set up by the lecturer to begin with as an appropriate forum.

It was clear that the increase in students' skills and understanding had led some instructors to raise standards. Interestingly, this phenomenon was most noticeable in the case of the instructor who reported his initial reluctance to adopt OASIS:

> [the students] do get very proficient at doing these things. We then have a tendency to shift the goalposts and that's a problem. In fact it's hard to think of

> a lecturer who hasn't shifted the goal posts… At the end of the course, he [a certain lecturer] would give them a test… he'd give them one problem to do and he'd get a pass rate that was about 30% to 50% and he always lamented it and said that the students are as dumb as two clucks and they won't do any work. Last year for the first time he was actually dragged kicking and screaming to get a suite of OASIS questions which were of the type and degree of difficulty that he wanted and he saw how good the students were at doing these, and his final test, which was in the same amount of time, was four of the questions that in previous years he's only asked one of, and more than 80% of the class got all four perfectly correct… the test was four times longer and they did three times as well. His attitude was, well, he's going to make the exam question a bit more challenging. Now I don't think that's what you should do.

In my courses I had also noticed a clear improvement in students' skill levels and understanding following the implementation of OASIS. However, in my courses the number of assessments had increased: the traditional written tests were augmented with OASIS assignments. Therefore the improvement could have been caused by the increase in the number of assessments, rather than by OASIS in particular. Of course, it was the automated nature of OASIS that had made it possible to increase the number of assessments and the amount of feedback to students. In some courses the implementation of OASIS had not increased the total number of assessments: the traditional tests had simply become half OASIS and half paper-and-pencil tests. In spite of this, an increase in student skill levels and understanding was still reported in all such courses. Instructors in such courses also reported a reduction in workload due to the reduced requirement for tests to be marked manually:

> [OASIS] lowered my marking load so that I could mark that test [half OASIS, half paper-and-pencil] in a third of the time.

Instructors considered that the initial implementation did require considerable time but, once they had put in this 'start-up' time and produced a set of OASIS questions, there were subsequent workload reductions. This reduction was for both skills practice and assessment:

> In terms of saving staff time, it [OASIS] does seem to cost a little bit extra the first time you use it because it's a bit of a learning curve in setting up some nice questions but once you use it the second time round you've already got the questions and it saves a lot of time especially in terms of paying markers and that kind of thing. I think even just replacing current small assignments with OASIS assignments saves quite a lot of time without any downside.

> I think [work load] goes down basically because one doesn't have to then plan how you're going to try and give that development of skill experience to the students. Essentially OASIS is doing that for you. Once you develop a set of questions that work reasonably well and also debug them, and that's actually

> not a trivial exercise, then you've basically got it for year after year and you've got a reasonably consistent platform there as well from which you can compare results from year to year.

In my case, the implementation of OASIS had not provided me with any workload reduction because any potential time-savings had simply been used to provide more learning and assessment opportunities for the students. One of my colleagues remarked that there was a happy medium where the students could be better served with OASIS and staff workloads could also be reduced:

> I think in terms of using [OASIS] to do new things we couldn't do before… it does cost a little bit of extra time but then we're doing a lot more that wasn't done before, students do appreciate that… I think the idea is to save a bit of time that can then be spent doing more.

In courses where OASIS assignments were used, both students and staff were concerned about students who relied on the efforts of others. These interviews revealed that, even in courses where OASIS was used only for supervised tests, staff were still concerned about students who relied on the efforts of others. When the test questions were chosen from the practice questions, one student could work out how to do a question and tell other students the formula or method. These students would then memorise the method or formula and use it in the test, without any real understanding, perhaps scoring a higher test mark than deserved. However, their learning would suffer and, in the long term, they would be disadvantaged:

> The only way I can see they can get value out of OASIS is… if the effort is 100% theirs. If they use somebody else's derivation, obviously they've watered down the experience so they're only doing it by rote.
>
> If they have done the problems by themselves, then they will certainly learn a lot. If, however, they do not do the problems by themselves, then their learning is compromised.

Students who did rely on the efforts of others as above could have been thwarted in tests by test questions that were similar but not identical to the practice questions. It should be noted, for traditional tests too, sometimes students prepare for certain questions in a mechanical way with little understanding. The instructors interviewed were certainly not against collaboration in general, provided all the students were making a genuine contribution to the group effort:

> The students sit around and get a friend to help them do the assignment. We're still better than we were with paper-based assignments which have worked for

> decades so I don't know that that's necessarily a problem, although I think we could be a little bit clearer about what our goals are with OASIS
>
> It's probably a good thing [for students] to help each other as long as there's not one free-loader who sits there and everyone does the question for them but I think students are quite good at not helping those people… I'm not sure if that's a really big problem… the key is not to make the unsupervised OASIS assignment worth more than a paper-based assignment.
>
> You go into a lab room, there's a new set of problems that none of them have ever seen and they'll sit round in a group of six or five and they'll talk about these and when they've talked about a particular problem they'll all scatter and go to the computer that they're logged on at and try that problem and if they can't get it right they'll come back and talk with the other guys about 'what method did you use?' so they actually do explore methods and techniques. I think that's wonderful. One of them even confessed that he thought it was cheating a bit. Well, we can do with more of that sort of cheating.

When staff were asked to voice their thoughts about the future of OASIS, a number stated that they might make use of a greater variety of questions and use OASIS more extensively. It was acknowledged that OASIS provided some opportunities that had not been realized as yet:

> I'll try [OASIS] in some of my own programming courses, that's a strength that hasn't been used yet. It's more a matter of trying it out and seeing what's possible.
>
> I think just the types of questions OASIS can ask, there's a huge range and the possibilities there are massive and we've only really scratched the surface on that. The system as it is, without any changes, is capable of a lot more than we've ever used. So I'm hoping to explore that myself next semester. There's a lot more room there.

Some staff also judged that OASIS could be used much more extensively, perhaps even to provide the majority of the assessment in courses, rather than the 10% to 15% it was typically providing:

> I think we should use it more intensively in our courses, rather than just use it as a little check system at the end and say 'oh yeah, they can do that'.
>
> I personally would like to see a different marking stratagem put in place where students' test pieces through the year might actually add up to 50% or even 60% of the final mark and in the final exam we eliminate all calculations altogether and [focus on] things which are really fundamental and really hard to do and really test their understanding.

There was no enthusiasm voiced for packaging OASIS with other software, although it was considered that it could sensibly be used in courses in conjunction with other software packages:

CS: Do you see it as being combined with some other software package like Cecil or LON-CAPA?

I: I think it's already designed to do one thing and do it well: it's designed to present questions and mark them… I think the future lies in a collection of applications that do one thing really well but which can talk to each other… it should be stand-alone but interface well with other things, not become part of them.

CS: I can see there's a subtle difference there.

I: if you have to choose one package that does lots of things you might get bits you don't like, whereas, if you can choose individual components and then put them together, you can choose bits that work best for you. Plus it makes it easier for development because you don't have to worry about the other components, you can go ahead and make big changes to improve something without worrying about the effect on other systems.

For the future development of OASIS, it was noted that, although the number of students using OASIS was increasing significantly, the number of staff using OASIS was growing less than might have been expected. Barriers to new staff adopting OASIS included the increased initial workload involved in producing OASIS questions and, possibly, a perceived lack of support:

> It's a different way of thinking so I think that the people who try out new things already have and are using it whereas the remainder of the people might not have time to try it or something…I think the software itself is okay but [to get more staff involved] I think there needs to be a system around it or people who are there to help… I think you need a human group to help people out at a certain time to show how to do things. That would be the next way to spread it further.
>
> To use OASIS, people have to spend a little bit more time than they otherwise would, especially to get started, and people are really pressed for time. Just being able to have someone who can help them [would be good]… helping them design questions in a way that would work well as a question, not just technically.

I had been aware of the initialisation difficulties with OASIS since my first involvement. It was hoped, in part, to address this issue with the creation of a software package that would assist instructors to create OASIS questions themselves. This package is discussed in more detail in the following subsection.

9.7 Reflections and plans for future action

This chapter is not followed by a chapter outlining the events of 2006. Although the development of OASIS and its implementation continued throughout 2006, this

account has to stop somewhere, and it does so at the end of 2005. Nonetheless, this subsection outlines both the reflections on the events of 2005 and those plans for future action that naturally arose from them.

The two action-research cycles of 2005 were very similar to those of 2004. Again, students made good use of OASIS and produced generally pleasing results in the examination questions supported by OASIS. As discussed in Subsection 9.3.4, one of the year-one questions had produced a lower-than-expected result, perhaps the result of a less-able student intake. The year-two results had been particularly pleasing, with pass rates on my two examination questions of 81% and 95%, respectively. Pass-rates for the other four examination questions ranged from 72% to 33%.

Unfortunately, mirroring the events of 2004, there had again been a problem with the first year-one OASIS assignment in 2005. This problem was caused by a software revision that had introduced a fault in the software. The 2004 problem had been caused by a combination of software, hardware and network systems that were inadequate to handle a particularly large load. If any lessons could be learned from these time-consuming and damaging problems, they were that careful forward planning is necessary for large-scale operations and that changes must be tested meticulously and then carefully introduced.

The introduction, in the year-one class, of assignment questions that were both previously unseen and different for different students appeared successful. The move to alternative questions had not produced significantly reduced assignment marks (see Subsection 9.3.2), although it would have largely eliminated one undesirable practice: students could no longer watch students sit the assignment earlier in the day, learn what questions were in the assignment, and practise those. The fact that this change had not reduced marks significantly suggested that the student practice outlined above had not been as widespread as some students reported. However, students still could, and sometimes did, work together in informal or formal groups and help each other complete the assignment. While this was judged to be a relatively minor problem in terms of mark-inflation, it was perceived by some students as a major problem. In particular, a significant number of students regarded the assignments as unfair because students could score highly through the efforts of

others. I did not want to move from assignments to tests for the year-one class because of the huge organisational problems involved in arranging computers and supervision for 600 students. However, I did want to make the assignments fairer and seen to be fairer. To this end, I decided to limit the extent that OASIS assignment marks could lift a student's overall course mark. Such a limit had been suggested by students during a focus-group discussion (see Subsection 8.2.3). I judged that such a limit was not needed for the year-two course. The following statement was added to the 2006 day-one handout for the year-one course:

> Please note that your OASIS assignment mark cannot improve your final mark for this course by more than 3% in comparison with your performance in the test and examination components.

This constraint would only affect students who scored far more highly on the OASIS assignments than they did on the other course assessments: two tests and one examination. For example, a student who scored 100% (i.e. 10/10) on the OASIS assignments and 70% (i.e. 63/90) on the other course components would be unaffected by the new constraint and would still receive a course mark of 73% (73/100). However, a student who scored 100% (i.e. 10/10) on the OASIS assignments and 60% (i.e. 54/90) on the other course components would receive a course mark of 63% (63/100) with the new constraint. This student would be penalised by just 1%: without the new constraint their mark would have been 64%. It was judged that such a penalty would not be seen as too punitive: the above examples show that a mark of 70% on the tests and the examination would enable a student to receive full credit for a mark of up to 100% on the OASIS assignments, whereas a mark of 63% on the tests and the examination would enable a student to receive full credit for a mark of up to 90% on the OASIS assignments. As yet this system has not been implemented in any courses in our department. However, it has been implemented successfully in other Engineering departments.

Software development had been continuous throughout this research programme. A student was hired over the 2004-2005 summer, funded by my research fund and by the Department of Physics, to improve the statistical functionality of OASIS. In some cases this entailed presenting data in a more graphical nature; in others, presenting new data that had not been presented, or even collected, before: for example, data on how long students took to complete assignments. During the 2005 academic year a

few changes were made to OASIS, most notably the facility to provide different assignment questions to different students. A perceived problem with the browser back-button during assessments was resolved for 2006: OASIS now informs students who use the back-button and make a second submission that they have already submitted answers and that they will be scored on their first submission. The software also captures all submissions and indicates on the assignment results page for staff those students who have attempted to submit more than one set of answers.

During the 2006 academic year two year-four students pursued a research project that involved improving the management and presentation of some of OASIS statistical data. For example, one part of the project aimed to present the average time students spent on each question. To make this meaningful, very short and very long times should be discounted when calculating the average. Some investigation was necessary to determine suitable lower and upper limits for the times. Another part of the project involved a "How am I doing?" link that students could access to find out how they had performed on each question: how often they had done each question, what their success-rate was, and when they had last attempted the question.

The interviews conducted with staff confirmed there were significant initial barriers for staff who wished to use OASIS. One of these was the fact that a large pool of questions needed to be created. Staff who did use OASIS in their courses had typically produced large numbers of questions initially; subsequently, however, they seldom added questions on a regular basis. I was the only instructor who had produced extra pools of questions that were used only for assessments. Such a pool would have eliminated the problem spoken of by a number of my colleagues: students preparing for OASIS tests by memorising how to do every practice question. These shortcomings were largely a result of the difficulty of producing OASIS questions. The process, which involved hiring students to encode the questions, was also time-consuming, and questions needed to be specified well in advance of when they would be needed.

In order to address this problem, the OASIS Committee applied for and received a Teaching Improvement Grant. This grant was used to employ a student over the summer of 2005-2006 to produce an OASIS question composer. This software

package would enable staff to create OASIS questions and upload them to the OASIS question database without any specialist programming skills. Staff would, however, need to specify the answers in terms of the question variables and produce images of circuit diagrams where relevant. The range for each variable would also need to be specified. Once this was done the software would automatically restrict the variable values to standard values within that range. For example, resistances would be restricted to the E24 series of values so that the resistance values in OASIS questions would be the same as those encountered in everyday practice. A screen-shot of the OASIS question composer is shown in Appendix five, Figure A-5-8.

In spite of the problem with the first year-one OASIS assignment I considered that, overall, the two action-research cycles of 2005 had been most successful. In terms of outcome validity, there was again a good body of evidence from examination results and interviews with both staff and students to suggest that student learning had been enhanced. This evidence was consistent with the evidence from the previous years. The OASIS student survey was not used in 2005, for the reasons advanced in Subsection 8.2.5. However, I considered that the lack of evidence from surveys was more than balanced by the evidence obtained from sources not formally accessed in 2004: postgraduates and staff. Further, the evidence of student learning provided by assignments in 2005 was somewhat more reliable than the 2004 evidence, given the use of alternative questions in Semester two, 2005.

Finally, the breadth and depth of the evidence from the implementation of OASIS in the Department of Physics was considerably greater in 2005 than in 2004, and, in itself, provided compelling justification for their continued use of OASIS. It was encouraging to note that the Department of Physics, in spite of losing their key OASIS proponent, was planning on increasing their usage of OASIS for 2006. In particular, they intended to move from four to 10 OASIS assignments each semester.

I personally had still not achieved the desired outcome of staff workload reduction. Rather, I had simply capitalised on the efficiencies created by the implementation of OASIS to deliver more practice opportunities and assignments to the students. However, providing these without the help of OASIS or similar software would have entailed an unmanageable workload. In courses where some activities had been

migrated to OASIS, there had been a clear workload reduction. This was certainly true for the large year-one Physics classes where OASIS assignments had replaced written assignments, and also for those ECE courses where written tests had been partly replaced with online tests.

In terms of process validity, I judged the rigour of the research at least as high as in previous years. As was usual, considerable informal feedback was received from a variety of sources. Formal feedback from student interviews was again collected and analysed. Following the 2005 teaching year, staff were also interviewed. Feedback was also gathered from postgraduates, mainly by email. Compared to the year-one and year-two students, both staff and postgraduates could provide a longer-term view of the implementation of OASIS. Except where the software itself or the nature of the implementation had changed, there were no discrepancies between these longer-term views and the views of the undergraduates. While the survey was not used in 2005, it was considered that the interviews with staff and the feedback from postgraduates provided adequate compensation and triangulation, particularly given that the longer-term perspective of the staff and the postgraduates agreed well with the shorter-term perspective of the undergraduates.

Observation, method and interpretation triangulation were all further increased by the involvement of the University's Department of Physics: having implemented OASIS with their large year-one classes in both 2004 and 2005, they were now well placed to provide a significant level of feedback about the effect OASIS had had on staff workload and student learning. The feedback from this department was certainly positive on both counts: in their judgement OASIS had radically reduced workloads while promoting student learning (Kruhlak, Mullins, Coghill, & Ng, 2005b).

The democratic validity of the action research was maintained through informal communication channels with staff and students (both face-to-face and email) as well as the more formal channels provided by interviews, class representatives, the Staff-Student Liaison Committee and the OASIS Committee. The Department of Physics commitment to OASIS at year-one level also effectively increased the numbers of students and staff that were having some influence on the research programme.

The dialogic validity of the research certainly reached a new peak in 2005. Through conference attendance and research visits to universities, I took part in considerable discussion with peers about OASIS and similar software packages. I also had a paper published in a Special Issue, devoted to showcasing exemplars of Web-based instruction, of the top-rated journal in the field of electrical engineering education (Smaill, 2005). My paper was one of 24 chosen from 86 submitted manuscripts. In 2005, 'Team OASIS' was also honoured at the American Society for Engineering Education / Australasian Association for Engineering Education 'Global Colloquium on Engineering Education' in Sydney, Australia with an award for Excellence in Engineering Education, Runner-up in the Curriculum Innovation section.

2005 marked the formal end of the action research programme. However, OASIS would still be used in 2006, as it had been in earlier years. The software would automatically collect data, and students and staff would still provide feedback about OASIS. Staff would analyse and reflect on this data and feedback, and would modify their practice accordingly. While the research would be less formal and exhaustive in nature, it would certainly continue. Further, with OASIS now relatively mature, and therefore requiring less effort to develop and maintain, possibilities existed for OASIS to provide a vehicle for new research: for example, research to determine the extent to which students already possessed, grasped, or retained certain concepts.

9.8 Summary

There were few major changes in the implementation of OASIS in 2005. For students, the most significant was the use of alternative questions in assignments in the year-one course 'Electrical Engineering Systems'. This change did not affect the assignment marks significantly, perhaps indicating that the problem of students determining the assignment questions earlier in the day and practising them prior to their own assessment was not as significant as some reports suggested. To partly counter continuing complaints about the practice of students working in groups on assignments, it was decided, for the year-one course in 2006, to limit the extent to which the OASIS assignment marks could lift the overall course mark.

Interview, email and informal feedback from undergraduates was augmented with feedback from staff and postgraduates. The Department of Physics, who had now used OASIS extensively for two years, also provided considerable feedback. There was very good agreement among all these sources. The instant feedback and repeatability of OASIS encouraged students to make good use of it for practice. There were reservations about OASIS assessments: students could help each other in assignments, and students could rote-learn how to do questions. Postgraduates, in particular, commented on the choice between surface and deep learning, perhaps because they had had more time to observe and reflect on the consequences of the choice. Staff frequently commented that they had adopted OASIS in order to cope with increasing class sizes and decreasing student work ethics. All staff considered that their decision to adopt OASIS had been vindicated.

Although the development of OASIS and its implementation continues, 2005 was the last year of the formal action-research programme. The following, final chapter offers an overview of the research, draws conclusions and presents some directions for the future.

Chapter 10. Conclusions

10.1 Purpose and outline of chapter

The purpose of this final chapter is to present a summary of the research programme and a synthesis of its outcomes and findings. Some directions for future development and research are also noted. An overview of the research programme is given in Section 10.2. The research findings are reviewed in Section 10.3. This review is in terms of the four research goals initially stated in Section 1.5. From the findings of the research it is possible to draw out some guidelines for the successful implementation of OASIS and, it is to be hoped, similar software packages. These guidelines are presented in Section 10.4. Issues concerning the quality and validity of the research are revisited in Section 10.5. Here the action research programme is evaluated according to the five criteria given by Herr and Anderson (2005). Some implications of the study and future directions for both the development and implementation of OASIS are outlined in Section 10.6. Future research possibilities, some focused on OASIS, others supported by OASIS, are also mentioned. Finally, a few concluding remarks are to be found in Section 10.7.

10.2 Overview of the research programme

The overall aim of this research was to develop, implement and validate a Web-based software package that, through providing practice and assessment opportunities, improves student learning and reduces marking and related mundane aspects of instructor workload.

The software package used was OASIS. This package, written and developed by ECE academics, already existed in prototype form when I joined the department in February, 2002, and the expertise was there to develop it further. Preliminary feedback (see Section 6.2) suggested that OASIS did enhance student learning and reduce staff workload. An examination of available software, coupled with a literature review, supported the decision to develop OASIS rather than implement an alternative, existing package.

Given my role as teacher, the traditional research ideal of observing without affecting the research environment would have been both impossible and unconscionable. An action research methodology was seen as best suited to my double role of teacher and researcher. This methodology enabled me, in the light of research findings, to modify the learning environment and enhance student learning. The action research proceeded through a spiral of one-semester cycles of planning, acting, observing and reflecting.

To maximize rigour, the research ran through eight cycles and involved considerable triangulation. OASIS itself collected much quantitative data. Further data was collected by interview, survey, and email, both formally and informally. This data came from three groups: current students, postgraduates, and ECE academics. These academics provided another source of observations and interpretations, as did the Physics Department academics, and an academic from CPD who was invited to interview the ECE academics and investigate the implementation of OASIS.

A literature review provided a sound foundation for the implementation and development of OASIS. This review had three main focus areas: assessment, feedback, and CBA. The findings of this review both validated and guided the development of OASIS. For example, the literature noted the great strength of the motivating power of assessment, and the pivotal role played by formative assessment. It also noted the importance of prompt feedback and, perhaps surprisingly, the fact that detailed feedback may be of no more benefit to learning than knowledge of the correct answer. Finally, it was noted that CBA can validly replace PPA provided certain conditions are met, such as being able to sight all questions and modify all answers prior to submission.

The overall research aim was expanded into four research goals. The extent to which these goals were met is discussed in the next section.

10.3 Review of the research findings

In this section the research findings are reviewed within the framework of the four research goals formulated at the start of the study. These goals were first presented in Section 1.5.

10.3.1 First research goal

The first research goal was:

> To produce a software package (OASIS) that meets certain criteria (e.g. Web-based, secure, robust, scalable, easy to use, provides prompt feedback, logs student activities and results) and can be used for student practice, assignments and tests in electrical engineering.

The associated research questions asked what criteria were appropriate for OASIS to meet, and whether it met them; and what student activities and results were appropriate for OASIS to record, and whether it recorded them.

To a large extent this goal has been achieved: for example, the software written to replace the prototype is Web-based, secure, robust, scalable, and provides prompt feedback. It is easy for students to use, as confirmed by surveys and interviews. It is regularly used for student practice and assessments in both electrical engineering and physics. The software also now records essentially *all* student activity; although there was a considerable period when it did not record the time students spent on each question. However, the first research goal has not yet been fully achieved. The focus in developing OASIS had been on students, and staff did not always find OASIS easy to use. For example, producing new questions and adding them to the OASIS question database was difficult, although it is hoped that the recently-created OASIS question-composer will make this easier. Further, while OASIS does log all activity, some statistical information has been difficult to extract. To be fair, the information available to instructors now appears generally adequate. It is in my capacity as a researcher that I need more statistical information. For example, I cannot currently see which actual questions an individual student has practised.

Instructors can see overall day-by-day usage patterns for individual students and for whole classes. Information about individual questions is also available: the number of students who have attempted a question, the total number of attempts, the mean

score and the standard deviation. This information is provided for practice questions and for each question used in every assessment. The live assessment information for staff is particularly good, with staff being able to monitor assessments as they proceed, and view individual student questions and answers, start times, and durations. The information can be sorted by name, start time, or mark rank (either total or individual question mark). In addition to accessing information, OASIS now makes it easy for staff to create tests and assignments, and to reset students' login passwords. Appendix five contains screen-shots illustrating some of these features.

The first research goal was intentionally somewhat imprecise: it did not attempt to specify all required features in precise detail from the outset of the research programme (although further detail was added after one year - see Section 6.5). There were two reasons for this. First, it was expected that these would be a matter for ongoing debate, and this certainly proved to be so. As one example, students often made a plea for model answers, or at least hints, while I wanted to provide answers only. There were pragmatic reasons for this: for example, providing the extra material would have required considerable staff time. There were also academic reasons: for example, model solutions are sometimes seen as the only solutions. Further, there is a good body of research (described in Section 3.5) that suggests detailed feedback beyond knowledge of the correct answer provides no extra benefit to learning. Currently, in almost all courses, the feedback consists of only the correct answer. In a few courses, some questions do provide simple hints.

Second, decisions about desirable features have been taken in tandem with the development of OASIS and its implementation. Features that were initially regarded as important became less so with time. For example, to stop students programming formulae into their calculators, a software calculator that students could use instead of their own calculators during assessments was initially deemed important. This became less important when previously-unseen questions were used in assessments, since the students could not determine in advance formulae for such questions. Conversely, features that were hardly considered early in the programme, such as a question-composer, a backup server and software documentation, become more critical as the circle of OASIS users widened.

In summary, this goal has been largely achieved. However, in all probability it can never be completely achieved. As more features are added, new possibilities open up and this may lead to demands for still further features. Also, as new users adopt OASIS, they bring with them their own unique demands.

10.3.2 Second research goal

The second research goal was:

> To incorporate features in OASIS that encourage staff to adopt it for their use with courses they teach. In particular, use of OASIS should reduce staff workload. This goal also includes the longer-term goal of extending the use of OASIS beyond the department, faculty and University.

The associated research questions focused on the instructors using OASIS and asked why they decided to use it, what they liked and disliked about OASIS, and what effect OASIS had on their workloads. Considerable informal feedback relevant to these questions was received on a regular basis. This informal feedback was supplemented with feedback derived from the interviews conducted with instructors, as described in Sections 7.7 and 9.6.

This goal has been partially achieved. Initially OASIS was only used in fractions of a few ECE courses. Now it is used in over a dozen courses (see Figure A-5-1), and in some of these the coverage is extensive. Initially all OASIS assessments were tests; now, although the number of tests has gone up rather than down, the bulk of OASIS assessment is via assignment. OASIS has also been implemented in some very large year-one classes, so the number of students using OASIS has grown out of proportion to the number of classes using it. This year-one implementation has been in both ECE and in the Department of Physics, and there is also a small level of use in the Department of Mechanical Engineering.

In some cases, the implementation of OASIS reduced staff workload, in other cases it did not. This depended on the nature of the implementation. In courses where OASIS replaced traditional assessments, there certainly was a workload reduction. This happened in all the physics courses and in some ECE courses. The largest savings occurred in the largest physics classes: replacing four written assignments with OASIS assignments in a class of 500 students saved literally hundreds of hours of marking time. Some ECE courses, notably at year-three level, replaced written tests

with tests that were half written and half computerized. This was also reported as saving considerable marking time: in fact the marking time was described by one academic as being reduced by two thirds (see Section 9.6).

In other courses, including the three in which I was involved, OASIS was used to increase the number of assessments. The instructors in these courses judged that the assessment schedule had been too light, particularly where formative assessment was concerned. However, the level of assessment had been limited by workload considerations. OASIS enabled further assessments to be added with little workload increase. In these courses, the motivation for using OASIS had not been reduced workload but the enhancement of student learning. It was judged that the learning enhancement was considerable and well worth the small increase in workload.

The way the second goal had been written, at the start of the research programme, suggested that if OASIS incorporated the right features and could be seen to reduce staff workload then it would be adopted by staff. However, interviews with instructors indicated that they had generally adopted it, not for workload reasons as such, but for the purpose of enhancing student learning (see Section 9.6). Instructors often spoke of how student engagement and performance had reduced over time. Sometimes this was a consequence of lower entry standards, sometimes a consequence of increased student numbers that forced changes in the way students were taught. OASIS was seen as providing a way to encourage students to engage with the course material and consequently lift their achievement levels. For example, OASIS could provide an assessment and feedback to students in week two or three of a semester; this was seen as much more useful and motivating than an assessment in week five or six that provided feedback two or even three weeks later. Student feedback consistently corroborated this belief; see, for example, Subsections 9.2.2 and 9.3.3. Instructors had a vision of what was an appropriate assessment regime for their courses but had not been able to achieve it because the workload involved would have been excessive. OASIS made their visions more achievable.

The difficulty of creating and adding new questions to OASIS was certainly a barrier to staff who wished to introduce OASIS to their courses. Apart from this difficulty, no barriers were explicitly mentioned in interviews and informal discussions with

staff. It is hoped the new question composer will help reduce this barrier. Of course, there is a natural resistance to change, and to explore a possible change takes time, time which is not always available to those whose schedule is already busy.

In summary, this goal has been partially achieved. OASIS has become a mainstream tool, with significant use beyond ECE; however, many ECE academics still do not use it. It is hoped the new question composer will reduce barriers to adoption. In some courses OASIS has reduced workload; in others OASIS has enabled increased assessment, particularly formative assessment. Staff who used OASIS to increase assessment did so to increase student engagement and learning. This, rather than reduced workload, was the main reason staff implemented OASIS.

10.3.3 Third research goal

The third research goal was:

> To implement the use of OASIS in such a way that student learning is enhanced.

The associated research questions focused on the ways students and instructors used OASIS, what students liked and disliked about OASIS, and the staff implementation and student usage patterns that correlated with increased student learning.

Not only was this a research goal, it was also one of my key goals as an instructor. Further, since the ability of OASIS to enhance student learning was the main motivation for most instructors to use OASIS (see Section 9.6), achieving the second research goal was largely contingent on achieving this goal. Therefore, once the software itself reached some level of maturity and its development slowed somewhat, this research goal assumed primary importance. Consequently, the associated research questions were visited regularly in a variety of ways. Student interviews were used in six of the eight research cycles. These are described in Sections 7.7 and 9.5, and Subsections 8.2.3, 8.3.5, 9.2.2 and 9.3.3. Student surveys dedicated to OASIS were used in three research cycles. These are described in Subsections 7.4.3, 7.6.3 and 8.2.4. Informal feedback from students was received on a regular basis via email, conversation and course survey forms. Two sets of interviews were also conducted with ECE academics, as discussed in Sections 7.8 and 9.6. Further input was sought from postgraduates (see Section 9.5) and the

implementation of OASIS in the Department of Physics was investigated too (see Section 9.4). I also considered examination results for my own courses in order to confirm the feedback from my students and instructor feedback about their own courses (see Subsections 6.4.5, 7.4.4, 7.6.4, 8.2.5, 8.3.6 and 9.3.4, and Section 9.7).

This goal was achieved: the implementation of OASIS did enhance student learning. However, the goal is imprecise, and the measurement of student learning was imprecise too, raising some issues about the validity of the conclusion. For the reasons explained in Sections 1.6 and 5.2, I had implemented OASIS with all my students, so I could not compare a 'control group' with an 'OASIS group'. Instead, comparisons were made between the results gained by current students (who used OASIS) and past students (who did not use OASIS), or between the marks current students scored on OASIS-supported questions and non-OASIS-supported questions in examinations. Clearly there are major issues with such comparisons. For example, students may have scored more highly on the OASIS-supported questions because those questions were easier, or the current students who scored more highly might have done so because they were more highly achieving in general. I addressed the first of these issues at the time by having my questions carefully checked by colleagues who assured me that my questions were of appropriate difficulty; in fact often my colleagues considered that my questions were too difficult. For the second issue, it was clear by virtue of reducing entry standards that the current students were on average less rather than more highly achieving. The comparisons were also made more tenable by the fact that the differences were quite large. For example, in the year-one course 'Electrical Engineering Systems' the failure rates for the questions supported by OASIS in the 2002 examination were less than half the failure rates for the corresponding questions (not supported by OASIS) in the 2001 examination. More details are given in Subsection 6.4.5. The comparisons between 2001 and 2002 were made more problematical because the lecturer had changed: I lectured in 2002 but not in 2001. Fortunately the improvements apparently due to the implementation of OASIS in my courses were paralleled by improvements observed by other instructors in their courses. These instructors were quite definite in their statements that student learning had been improved, and they had, in almost in every case, taught the course both before and after the introduction of OASIS. Furthermore, they had made these statements to both me and an academic from CPD who had

conducted staff interviews a few years prior to my own interviews. These interviews are described in Sections 7.8 and 9.6. It was also encouraging that no staff member who had started using OASIS had later stopped using it. Staff observations were backed by corroboration from the students themselves. Surveys, emails and interviews all gave the same consistent and clear message: student learning had been enhanced. More will be said about the validity of the research in Section 10.5

Students were also clear and in good agreement about their likes and dislikes. They appreciated the instant feedback, the way questions could be repeated with different numerical values, and the fact that OASIS was easy to use and could be used at both home and university.

Students often expressed a wish that OASIS provided hints or model answers, although some students argued against this, stating that it was unrealistic and that students could become over-reliant on such things and fail to develop the important skill of choosing appropriate methods (See Subsection 8.2.3). Another very commonly-expressed wish was that OASIS covered more topics and more courses.

While students were enthusiastic about the use of OASIS for practice, they voiced strong reservations about its use for assignments (see, for example, Subsection 9.3.3). These reservations were centred on the fact that students could rely on the efforts of others and score well in assignments without having a good grasp of the material themselves. These reservations, much more commonly expressed by year-one than year-two students, led to regular improvements in the way I conducted assignments. Year-one assignments were first run in 2003, with all assignment questions being chosen from practice questions. Some students were unhappy with this, noting that some of their peers used spreadsheets prepared by others to do the assignments (see Section 7.7). In 2004, I used previously-unseen questions in assignments to render such spreadsheets useless. However, student feedback now shifted its focus to another concern: students could watch other students completing their assignments, observe what questions were in the assignment, practise just those questions and then sit the assignment themselves (see Subsection 8.3.5). To answer this criticism, in 2005 the software was improved to allow alternative questions in assignments. Different students now received different questions in their

assignments, not just different numerical values in the same questions. With this change, students could no longer prepare for an assignment by simply learning how to do several questions: the total number of questions involved was quite large and even determining all of them would have been a considerable task. Interestingly, with all these changes designed to eliminate 'excessive help', the assignment marks hardly changed at all (see Subsection 9.3.2). Further, students who sat their assignments early in the day tended to score more highly than those who sat later in the day. These facts suggested that the practices students had complained about were not widespread or, if they were, they had no great effect on results. However, the complaints did not stop. Now they focused on group efforts. Some groups were quite formal and had practised extensively: the members were able to do their own assignments and check the assignments of the other group members in the allocated time. Other groups were very informal, with students simply asking those who happened to be nearby for help if they ran into difficulties with an assignment question. Interviews and emails (see Subsection 9.3.3) showed students to be concerned about these activities: they wanted students to be judged on their own merits, not the merits of those around them while they sat the assignment. I was not overly concerned about groups that consisted of students who had practised well for assignments. However, I did not want to reward students who chose to rely on the efforts of others. To this end, for the year-one class of 2006, I introduced a rule, suggested originally by students during a focus-group discussion (see Subsection 8.2.3), that limited the extent to which OASIS assignment marks could lift the overall course mark. Details of the rule were given in Section 9.7.

I made these changes to the assignments for two reasons. I wanted the assignments to be as motivating as possible and as fair as possible. These two are connected: making the assignments fairer helped make them more motivating too. If students believe they can easily score highly on assignments by relying on the efforts of others, some are likely to do just that. Conversely, if they believe their own efforts are necessary and will be rewarded, they will increase their efforts. To make sure that effort was rewarded, I kept the assignment questions similar to the practice questions. Students were aware of these considerations, and they did state that the assignments were motivating. OASIS records also showed the dramatic increase in practice activity immediately prior to assignments (see, for example, Subsection 7.6.2).

The student concerns that led me to change the way I implemented assignments could have been defused by moving to supervised tests. However, there were two reasons why I did not do this. First, we did not have the capacity to manage 600 students for a computer-based, supervised assessment. Second, I wanted to have regular assessments, but regular tests would have been too stressful for the students. Courses that did have OASIS tests had fewer assessments, perhaps three tests in total. In these courses, because the assessments were supervised, there had been little need to change the way the assessments were conducted. Usually, the tests comprised questions that the students had practiced on previously and all students received the same questions, albeit with numerically different variations.

Academics used OASIS in a considerable variety of ways. All of them were reported as increasing student learning. Even simply providing OASIS for practice without the motivation of OASIS assessments led to increased student learning. Of those instructors who used OASIS, some used it for assignments, others for tests. Some of the tests were entirely OASIS, others were half OASIS and half paper-based. However, there were some points of consistency. First, instructors had adopted OASIS assessments to motivate students to engage more with the course material. This engagement was largely via OASIS practice questions. Assessment questions were similar or identical to these. Second, instructors recognized that OASIS could not assess all that they wanted to assess. Typically only 10% of the total course marks were allocated to OASIS assessments. The OASIS assessments were not viewed as highly important in their own right; more important was the fact that they motivated students to engage with the practice questions: this engagement with the practice questions was a critical element in student learning and skills development.

There was even more variety in the ways students reacted to OASIS. A small minority of students did not use it for practice at all, and some of these also failed to submit assignments. Other students used it extensively. While some students predominantly interacted with it online, others printed out the practice questions and spent only a small fraction of their time online, preferring to write out notes or solutions to the questions on paper. In general, student engagement with OASIS increased dramatically just before assessments (as illustrated by Figures 7.6.2 and

9.4.2); however some students did start to work through the practice questions well before the assessment date, returning to review them just prior to the assessment, while others left virtually all their efforts to the last day or two. Student choices in this regard were highly influenced by workload demands in other courses. When workloads were low, most students preferred to start work on the practice questions as soon as they learned the skills to do so. However, when faced with a number of deadlines, students tended to handle tasks serially. In such cases, work on the practice questions often began in earnest only a day or two before the assessment.

Students adopted both surface and deep learning approaches to the OASIS practice questions. Some students were conscious of the choice; stating that they aimed to understand as much as possible, but that time pressures could lead them to adopt a mechanistic approach to their work. Students also reported that, when asking for help with questions, their peers sometimes gave them a series of steps that they could not justify at a deep level. Postgraduates, as reported in Section 9.5, were particularly aware of the choice to be made between surface and deep approaches: they had repeatedly seen the negative consequences for students who chose surface learning.

Students also displayed a large variety of approaches to sitting OASIS assignments. Assignments were sat at home, at university, and occasionally in Internet cafes. Some students sat assignments on their own, others in groups. Some of these groups were quite formal, with each member checking the work of all other members, or with half the group sitting the assignment, while the other half watched and checked. Informal groups also formed when a number of students were in a computer-lab at the same time and some of those sitting the assignment asked others for help.

Given the wide variety of approaches to OASIS practice and assessment outlined above, it was not possible to pinpoint a particular student usage pattern that correlated with increased student learning. This was true even for assignments: while it might seem reasonable that students who planned to sit the assignments on their own would be more motivated and therefore make more progress in their learning than students who planned to sit as one of a group, it was clear that many of the latter students in fact worked very hard to ensure mastery of the material. Perhaps being in a group was motivating in itself. Conversely, there were many examples of students

who sat the assignments individually and who scored poorly through lack of practice. Sitting assignments individually or in groups did not in itself greatly determine the level of student learning.

For OASIS practice it was not possible to identify a particular student usage pattern that correlated with increased student learning. While adding OASIS practice to courses (with or without OASIS assessments) definitely boosted achievement, there was a surprisingly weak positive correlation between the number of OASIS answers students submitted in practice and their examination marks. This issue is discussed in Subsection 6.4.3, where it was suggested that mindful engagement with the material was more important, and that engagement could not be measured by number of answer submissions, or even time elapsed between sighting questions and submitting answers. Students could also print out questions and engage with them offline, or decide that they would put little time into OASIS and instead put more time into the course notes, the textbook, or the tutorial problems. These activities could compensate for reduced online OASIS activity.

Both instructors and students accepted that it was, in general, good to start tasks well before deadlines, allowing more time to grasp concepts and develop skills. However, students frequently did not take their own advice on this matter. The OASIS practice records showed remarkable increases in activity (online activity, at least) shortly before assessments. Usually this was simply a result of procrastination or of the pressure of other demands on their time. Sometimes, however, it was a strategic decision: students looked over the practice questions and decided that they could be mastered in a few days and that the most benefit would be gained by doing them as near as possible to the assessment.

In summary, the goal of enhancing student learning through the implementation of OASIS was achieved. However, the extent of the enhancement could not be precisely measured. There was considerable variety in usage patterns. Students could use OASIS in many different ways and still enhance their learning. Similarly, how OASIS was implemented in courses varied from course to course. In each case staff judged that their implementation was successful. The one common ingredient was

the provision of OASIS practice questions, almost always motivated by OASIS assessments.

10.3.4 Fourth research goal

The fourth research goal was:

> To identify best-practice strategies for the implementation and use of OASIS.

The associated research questions asked how OASIS should be implemented and used to best promote student learning. The third research goal and its associated research questions were intended to foreshadow the fourth goal and provide the answers to its associated research questions. However, as the previous subsection makes clear, no single best-practice strategy emerged, for either staff or students. Perhaps it was unrealistic to expect such an outcome. However, the research findings do enable the identification of some guidelines for the successful implementation of OASIS. These are presented in the next section.

10.4 Guidelines for successful implementation of OASIS

While the following guidelines are derived from the findings of this research, it is hoped that they may have a more broad applicability. As noted above, the one common ingredient in successful implementations of OASIS was the provision of OASIS practice questions, almost always motivated by OASIS assessments. It is the practice questions rather than the assessments by themselves that are the key to increased student learning. As common sense suggests, the set of questions must be adequate in number, cover the requisite areas of learning adequately, and be graduated in difficulty. In particular, some very easy questions should preface the set in courses where students meet OASIS for the first time. OASIS should, of course, be complemented with a full range of learning resources. In particular, OASIS questions should not be the only ones students have for practice. Questions from textbooks and other sources should be provided too: although most students responded positively to OASIS, some preferred to work predominantly with other learning tools, even when preparing for OASIS assignments.

The instant feedback and the ability to repeat questions with different numerical values are key elements in making OASIS practice attractive to students. Students

often voiced a desire for hints or model answers, while others noted that this can be abused and can send the message that a particular question must be solved in a particular way. Guided by the research literature on extent of feedback (Section 3.5), OASIS was implemented almost entirely without hints. However, students will request help with some questions, so instructors must be available, either by email or in person, to provide this help promptly. Technical help must also be available to help students with log-in problems and similar.

Assessments are a key part in the implementation since they motivate students to practise, particularly when the assessment questions are similar to the practice questions. For supervised tests, instructors generally found it acceptable to use tests consisting of actual practice questions. Instructors need to be aware that students can programme formulae into their calculators and must decide whether such a practice is acceptable.

Assignments have the advantage that they can be scheduled more frequently than tests: they are less stressful and require less organisation on the day. For particularly large classes they may be the only viable assessment option. However, assignments have the disadvantage that they are open to cheating. This is both a perceived and an actual problem. In order to best address this, the assignment questions should include both previously unseen and alternative questions, so that different students receive different questions, not just numerically different versions of the same questions. Limiting the amount by which OASIS assignment marks can lift overall course marks should also be considered as an option.

Where OASIS tests are based on practice questions, the percentage of the course mark derived from the tests should be limited to about 10% or 15%. In courses where OASIS assignments are used, their contribution should be similarly limited. Students were happy with this level of contribution, seeing it as a valid reward for their practice, but they saw higher contributions as problematic, perhaps encouraging cheating or a reliance on rote learning.

Tests that comprise previously-unseen questions could certainly contribute more highly. Some instructors considered that such tests could contribute 50% or even

more to the final grade. Certainly, as noted in the literature review in Chapter 4, CBA in some institutions does contribute well beyond this level. However, CBA should only contribute at high levels to the final grade if adequate questions exist to assess the required range of skills, knowledge and understanding.

Often in university courses summative assessment is over-emphasised while formative assessment is under-emphasised, perhaps even virtually ignored. Both types of assessment are essential, and both can be provided very well by OASIS. In particular, CBA can provide prompt feedback without greatly increasing staff workload and so it is ideally suited to formative assessment. OASIS assignments can be given just two or three weeks into a course, providing significant and early feedback to students. Such assessments, aimed at informing students about their progress and achievement, should carry little weight in the final course grade. I regularly used OASIS assignments just two weeks into courses to provide feedback. Though these assignments contributed only 2% to the final course grade, students prepared extensively for them. Such low-stakes assignments can be used regularly to provide frequent, timely feedback without over-stressing students.

While OASIS is quite flexible in its application, my research findings suggest that following the above general principles will provide a sound basis for enhancing student learning.

10.5 Quality / validity of the research

Herr and Anderson (2005) identified five validity/quality criteria for action research. These criteria were described in Section 5.5. In this section I discuss the extent to which this study has met these five criteria.

10.5.1 Outcome validity

Outcome validity concerns the extent to which actions occur that solve problems or improve matters in the research setting. Herr and Anderson's (2005) associated action research goal is "the achievement of action-oriented outcomes" (p. 55). This criterion has been met. The software package OASIS was developed and subsequently achieved a good level of acceptance both inside and outside ECE. It is

now regarded as a mainstream learning tool. Through the use of OASIS, two key problems have been addressed with a good measure of success: high instructor workloads and lack of student engagement and learning. However, in courses where OASIS has been used to provide extra practice and assessments, student learning has been enhanced but instructor workload has not diminished.

10.5.2 Process validity

Herr and Anderson (2005) espouse a broad view of process validity. Their associated action research goal is "a sound and appropriate research methodology" (p. 55). To be both sound and appropriate, the research should, for example, involve triangulation and a series of reflective cycles in order to produce quality evidence, and also allow ongoing learning of those involved and promote quality relationships amongst participants. This criterion has largely been met. The requirements of conducting the research for a doctoral dissertation have somewhat restricted the extent of collaboration and the promotion of quality relationships amongst participants. However, apart from this restriction, the research methodology has been sound and appropriate. The choice of action research was certainly appropriate for an instructor-researcher. Section 5.5 discussed features of good action research. These features are now revisited.

Dick (1993) and Robson (2002) argue respectively for adequate iteration and prolonged involvement in the study. The present study involved eight main iterative cycles over four years. Triangulation is, of course, seen as a key procedure in establishing validity (Cardno, 2003; Costello, 2003; Dick, 1993; Herr & Anderson, 2005; Robson, 2002). Ideally this triangulation should be of methods, perceptions, observations and interpretations (Cardno, 2003). A number of different methods were involved, including interviews, surveys and quantitative analysis of OASIS data. Interviews and emails provided much feedback from undergraduates, postgraduates and instructors. These individuals brought their own perceptions, observations and interpretations to the study, increasing triangulation. The experiences of the instructors in the Department of Physics as they implemented OASIS provided further triangulation, as did the methods, observations and interpretations of the CPD academic who conducted interviews with ECE instructors.

The high level of agreement shown across the range of these observers is one of the strengths of this study.

A number of commentators note that the maintenance of an audit trail plays an important part in establishing validity (Cardno, 2003; Costello, 2003; Robson, 2002). All documentation has been retained, together with all survey forms. These items are stored in a locked filing cabinet in a locked office used solely by myself. All emails relevant to the study have been saved securely and backed up in electronic form, as have all recordings of interviews, together with their transcripts. Data collected by OASIS is stored on the OASIS server and is also securely backed up.

Critical appraisal of the findings and negative case analysis is also important for establishing validity (Cardno, 2003; Costello, 2003; Dick, 1993; Robson, 2002). Findings should be fed back to participants to check that the conclusions are trustworthy (Cardno, 2003). This was regularly done during interviews and focus-group discussions. My method here was to question students rather than present my conclusions and ask the students if they agreed. Negative cases also had an important role to play in this study, for they were seen as indicating ways in which OASIS and its implementation could be improved. In particular, negative student feedback did lead to a number of significant changes in the way OASIS assignments were conducted. These are discussed in Subsection 10.3.3. Student feedback also led me to look closely at the lack of hints and model answers provided by OASIS; however, after discussions with students, staff, and an extensive review of the relevant research literature, I decided to continue with the status quo.

10.5.3 Democratic validity

Democratic validity concerns the extent to which stake-holder perspectives and interests are taken into account. The research should produce change and knowledge that benefits the participants. Herr and Anderson's (2005) associated action research goal is "results that are relevant to the local setting" (p. 55).

As noted in Section 5.5, achieving democratic validity in this study essentially depended on achieving outcome validity and process validity. This is because the study aimed to benefit the stake-holders: the students and the instructors. Further, in

order to determine how best to deliver this benefit and verify that it had in fact been delivered, the viewpoints of students and instructors needed to be regularly sought.

Ideally, to achieve democratic validity, the research should have been fully collaborative. However, as noted in the previous subsection, the requirements of conducting the research for a doctoral dissertation made this impossible. Apart from this, democratic validity has been achieved because the study has met the requirements for outcome validity and process validity.

10.5.4 Catalytic validity

As noted in Section 5.5, a relatively mild interpretation of catalytic validity was used in this study; one based on action and increased participant understanding rather than reality transformations. Herr and Anderson's (2005) associated action research goal of "the education of both researcher and participants" (p. 55) is in agreement with such an interpretation. This interpretation reduces catalytic validity to elements of the previous three validity criteria. These have been dealt with already.

10.5.5 Dialogic validity

Dialogic validity concerns the quality of research as monitored by peer review. Herr and Anderson's (2005) associated action research goal is "the generation of new knowledge" (p. 55). This study has achieved dialogic validity through peer-reviewed publications, refereed presentations at conferences, discussions with like-minded researchers in New Zealand and overseas, and frequent discussions with fellow academics at UoA.

Appendix Nine lists relevant publications. Most of these are conference proceedings, and the associated conferences gave me good opportunities to present my findings and discuss them with my peers. While only two journal papers have been published to date, the paper published in the IEEE Transactions on Education is of special note. The IEEE is the world's largest body of professionals, and their journal is top-rated in electrical engineering education. The paper was published in a Special Issue devoted to showcasing exemplars of Web-based instruction. It was one of 24 chosen from a total of 86 submitted manuscripts.

Further confirmation came with an award I received on behalf of the 'OASIS Team' at the ASEE/AAEE 4th Global Colloquium on Engineering Education, Sydney, Australia, 2005. The award was for Excellence in Engineering Education, and the team was runner-up in the Curriculum Innovation category.

All five quality/validity criteria advanced by Herr and Anderson for action research have now been discussed, and it has been argued that the research programme has adequately met these criteria.

10.6 Implications and future directions

The OASIS software and its implementation were developed in tandem over a period of some years. This development was driven and determined by my research, by instructors and students, and by resource constraints. Favourable comments from instructors and students using OASIS, the increasing usage of OASIS, and this study itself, all attest that OASIS can achieve the goals of reducing staff workload and enhancing student learning. In general, this study is one voice of many affirming that CAL and CBA can achieve these goals. Given that the future is likely to bring more affordable and more powerful computers, networks and Internet facilities, and quite possibly even lower staff/student ratios, it would seem clear that there will be an increasing role for CAL and CBA to play in education at tertiary and other levels. While many of the findings from this research programme are specific to OASIS, a number have more general application. Thus, it is hoped that much of Section 10.4, 'Guidelines for successful implementation of OASIS', will be relevant to many instructors implementing CAL and CBA, whether or not they use OASIS.

Many of the findings of this study are likely to apply beyond CAL and CBA. For example, this research programme has produced a wealth of information about student study patterns, and the rapid increase in student activity just prior to OASIS assignments (illustrated well by Figure 7.6.2) is likely to be typical of student behaviour prior to assessments in general, not just computer-based assessments. Of course OASIS, with its automatic logging of data, provided a ready insight into student activity. Similarly, student interviews revealed some fascinating insights into strategies students adopted in approaching and tackling OASIS assignments; for

example: spreadsheet construction, a book for written solutions for every practice question, and various team-work approaches. It is likely that these or similar strategies are adopted by many students, not just those confronted with OASIS assessments. Further practical research, or a meta-study of existing research, could determine to what extent some of the student behaviours identified in this study are found elsewhere. While this study has answered some questions, it has raised others.

Certainly, there are a number of possibilities for future research. Some of these focus on OASIS itself. For example, while it is clear that providing OASIS in courses does enhance student learning, as yet no clear, strong mathematical correlation has been established between OASIS usage and examination results. Using the number of OASIS answers submitted as measure of OASIS usage produced only a weak correlation. Now that OASIS is able to collect data on time spent by students on questions, there are some new possibilities for exploration. For example, perhaps OASIS usage could be measured by the number of answers submitted to questions where the time between sighting the question and answering it was more than 30 seconds; or perhaps OASIS usage could be measured by the total time spent on questions (perhaps eliminating any times in excess of 45 minutes on one question).

OASIS can also facilitate research. For example, OASIS can deliver to students a set of questions designed to measure their grasp of certain concepts. This could be done at the start of a course and again at the end to check on student progress. The same set could be used the following semester or year to confirm retention. OASIS makes these scenarios possible since the setup and marking load is very low.

There are a number of possibilities for the future development of OASIS as a learning tool. For example, questions that require 'fuzzy marking' (see Section 6.5) are already supported by OASIS. However, instructors have not yet begun to use these. Most questions are not marked consequentially by OASIS, though the facility to do so exists. Few instructors in computer-systems engineering (one of the three degree programmes available in ECE) currently use OASIS, even though some have stated that OASIS could examine some important skill areas that are very hard to test by PPA because of the prohibitively long marking time.

Given enough OASIS questions, and enough computers, OASIS tests could contribute 50% or 60% to the final course grade in some courses. The rest of the grade would come from the final examination, where those skills not amenable to current OASIS testing would be assessed. This possibility was advanced by staff in interviews (see Section 9.6). The question-composer now makes it much easier for staff to construct OASIS questions: this should lower the barriers for new adopters. However, human support was noted as being a key ingredient for new adopters.

Looking beyond ECE, other departments have expressed an interest in OASIS. There is also some interest in using OASIS to remotely assess potential postgraduate students. Students could take an online entry test. The results would indicate to them whether they would be likely to pass a second, supervised test administered at UoA. Such an arrangement would hopefully eliminate students who were unsuitable before they travelled to UoA. Interest has also been expressed in OASIS by some high schools. One barrier to providing OASIS to secondary schools is the potentially very high server load. There are several dozen high schools in Auckland alone. However, it is possible that some commercial arrangement could be entered into, and some schools have indicated that this could be acceptable.

10.7 In conclusion

This study was intended to have both local and global significance. Locally, the aim was to develop and implement OASIS, thus improving student learning and reducing instructor workload. Globally, the aim was to add to the body of knowledge of CAL and CBA implementations, particularly where question types other than multi-choice are involved. Both these aims have been largely met. Locally, OASIS is accepted as enhancing student learning and thousands of students use it each year. Globally, this study adds its voice to the existing studies that speak in favour of CAL and CBA, often reinforcing what has already been said, sometimes saying something new.

However, while this dissertation ends here, the development and implementation of OASIS continues apace. Increasing computer power and affordability will bring possibilities well beyond those realised to date. We are merely at the start of exciting times.

References

Aizcorbe, A., Corrado, C., & Doms, M. (2000). *Constructing price and quantity indexes for high technology goods.* Paper presented at the National Bureau of Economic Research workshop on Price Measurement, Cambridge, Mass.

Albertelli, G., Kortemeyer, G., Sakharuk, A., & Kashy, E. (2003). *Personalized examinations in large on-campus classes.* Paper presented at the Frontiers in Education, Boulder, CO.

Altrichter, H., Posch, P., & Somekh, B. (1993). *Teachers investigate their work: An introduction to the methods of action research*. London: Routledge.

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261 - 271.

Anderson, A. (1996). Predictors of computer anxiety and performance in information systems. *Computers in human behavior, 12*(1), 61-77.

APA. (1986). *American Psychological Association on Professional Standards and Committee on Psychological Tests and Assessment: Guidelines for computerbased tests and interpretations*. Washington, DC: APA.

Arguinis, H. (1993). Action research and scientific method: presumed discrepancies and actual similarities. *Journal of Applied Behavioral Science, 29*(4), 416-431.

Ary, D., Jacobs, L. C., & Razavieh, A. (2002). *Introduction to research in education*. Belmont, CA: Wadsworth/Thomson Learning.

Baafi, E., & Boyd, M. (2001). *Presenting a first year engineering computing subject using WebCT.* Paper presented at the 12th Annual Conference, Australasian Association for Engineering Education, Brisbane, Australia.

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effects of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C.-L. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research, 85*(2), 89-99.

Bassey, M. (1998). Action research for improving educational practice. In R. Halsall (Ed.), *Teacher research and school improvement: Opening doors from the inside* (pp. 93-108). Buckingham: Open University Press.

Batchelor, R., & Jungic, V. (2004). Online interactive homework with LON-CAPA. from http://www.sfu.ca/lidc/intouch/intouch5.pdf

Beh, B. (2004). OASIS performance investigation: Unpublished part-four project report, University of Auckland, New Zealand.

Bell, B., & Cowie, B. (2001). *Formative assessment and science education*. Dordrecht, The Netherlands: Kluwer Academic.

Berry, F. C., DiPiazza, P. S., & Sauer, S. L. (2003). The future of electrical and computer engineering education. *IEEE Transactions on Education, 46*(4), 467-476.

Bigdeli, A., Boys, J. T., Calverley, P., & Coghill, C. (2001). "OASIS": A new web-based tutorial and assessment system, *12th Annual Conference, Australasian Association for Engineering Education*. Brisbane, Australia: Queensland University of Technology.

Biggs, J. (1999). *Teaching for quality learning at university*. Buckingham, UK & Philadelphia, PA: Society for Research into Higher Education & Open University Press.

Black, P. (1963). Examinations and the teaching of science. *Bulletin of the Institute of Physics and the Physical Society, 14*(8), pp. 202-208.

Black, P. (1998). *Testing: friend or foe? Theory and practice of assessment and testing.* London: Falmer.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education, 5*(1), 7 - 74.

Black, P., & Wiliam, D. (1998b). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Blackboard. (2006). From http://www.Blackboard.com

Blaxter, L., Hughes, C., & Tight, M. (1996). *How to research*. Philadelphia, PA: Open University Press.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Bonham, S., Beichner, R., & Deardorff, D. (2001). On-line homework: does it make a difference? *The Physics Teacher, 39*(5), 293-296.

Brooks, A., & Watkins, K. (Eds.). (1994). *The emerging power of action inquiry technologies*. San Fransisco, CA: Jossey-Bass.

Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.

Buchanan, T. (2000). The efficacy of a World Wide Web meidated formative assessment. *Journal of Computer Assisted Learning, 16*, 193-200.

Bull, J., & Collins, C. (2002). The use of computer-assisted assessment in engineering: some results from the CAA national survey conducted in 1999. *International Journal of Electrical Engineering Education, 39*(2), 91-99.

Bull, J., & McKenna, C. (2004). *Blueprint for computer-assisted assessment*. London: RoutledgeFalmer.

Campion, M. A., & Lord, R. G. (1982). A control systems conceptualization of the goal-setting and changing process. *Organizational Behaviour and Human Performance, 67*, 265-287.

Cardno, C. (2003). *Action research: A developmental approach*. Wellington: New Zealand Council for Educational Research.

Carr, W., & Kemmis, S. (1986). *Becoming critical: education, knowledge and action research.* Lewes, UK: Falmer.

Cecil. (2006). From http://www.cecil.edu

Charman, D. (1999). Issues and impacts of using computer-based assessments (CBAs) for formative assessment. In S. Brown, P. Race & J. Bull (Eds.), *Computer-assisted assessment in higher education* (pp. 85-93). London: Kogan Page.

Charman, D., & Elmes, A. (1998). Formative assessment in a basic geographical statistics module. In D. Charman & A. Elmes (Eds.), *Computer based assessmnet (Vol. 2): Case studies in science and computing*. University of Plymouth: SEED Publications.

Chen, P. M. (2004). An automated feedback system for computer organization projects. *IEEE Transactions on Education, 47*(2), 232-240.

Chou, C. (2000). Constructing a computer-assisted testing and evaulation system on the world wide web - the CATES experience. *IEEE Transactions on Education, 43*(3), 266-272.

Chua, S., Chen, D., & Wong, A. (1999). Computer anxiety and its correlates: a meta-analysis. *Computers in Human Behavior, 15*(5), 609-623.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education*. London: Routledge Falmer.

Connelly, F. M., & Clandinin, J. (1990). Stories of experience and narrative inquiry. *Educational Researcher, 19*(5), 2-14.

Conole, G., & Warburton, B. (2005). A review of computer-assisted assessment. *ALT-J, Research in Learning Technology, 13*(1), 17-31.

Costello, P. J. M. (2003). *Action research*. London: Continuum.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Cunningham, J. B. (1983). Gathering data in a changing organization. *Human Relations, 36*(5), 403-420.

Davidovic, A., Warren, J., & Trichina, E. (2003). Learning benefits of structural example-based adaptive tutoring systems. *IEEE Transactions on Education, 46*(2), 241-251.

Day, J. D., & Cordon, L. A. (1993). Static and dynamic measures of ability: an experimental comparison. *Journal of Educational Psychology, 85*, 76-82.

de Cock, C. (2005). Action research: In search of a new epistemology. In B. Cooke & J. W. Cox (Eds.), *Fundamentals of action research* (Vol. 3, pp. 371-375). London: Sage.

Deeks, A. (1999). *Web-based assignments in structural analysis.* Paper presented at the Australasian Association for Engineering Education 11th Annual Conference, Adelaide, Australia.

Dick, B. (1993). You want to do an action research thesis? Retrieved May 16, 2002, from http://www.scu.edu.au/schools/gcm/ar/art/arthesis.html

Dick, B. (1997). Action learning and action research. from http://www.scu.edu.au/schools/gcm/ar/arp/actlearn.html

Dick, B. (2002). Action research: action and research. from http://www.scu.edu.au/schools/gcm/ar/arp/aandr.html

Dill, D. D., & Sporn, B. (1995). The implications of a postindustrial environment for the university: An introduction. In D. D. Dill & B. Sporn (Eds.), *Emerging patterns of social demand and university reform: Through a glass darkly* (pp. 1-19). Oxford: Pergamon.

Elliott, J. (1991). *Action research for educational change*. Philadelphia, PA: Open University Press.

Elton, L. R. B., & Laurillard, D. M. (1979). Trends in research on student learning. *Studies in Higher Education, 4*, 87-102.

Entwistle, N., Hounsell, C., Macaulay, C., Situnayake, G., & Tait, H. (1989). *The performance of electrical engineers in Scottish education*. University of Edinburgh: Report to the Scottish Education Department, Centre for Research on Learning and Instruction, Department of Education.

Excell, P. S. (2000). Experiments in the use of multiple-choice examinations for electromagnetics-related topics. *IEEE Transactions on Education, 43*(3), 250-256.

Eysenck, M. (1988). Anxiety and attention. *Anxiety Research, 1*(1), 9-15.

Fernandez, G., John, S., & Netherwood, G. (2001). *Teaching engineering with an on-line student-centred environment.* Paper presented at the Australasian Association for Engineering Education 12th Annual Conference, Brisbane, Australia.

Fowell, S., Southgate, L., & Bligh, J. (1999). Evaluating assessment: the missing link? *Medical Education, 33*(4), 276-281.

Freilich, M. B. (1987). A student evaluation of teaching techniques. In M. S. Gupta (Ed.), *Teaching engineering: A beginner's guide*. New York: IEEE Press.

Fulkerson, F., & Martin, G. (1981). Effects of exam frequency on student performance, evaluations of instructor, and test anxiety. *Teaching of Psychology, 8*(2), 90-93.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction*. White Plains, NY: Longman.

Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: choosing and using diverse approaches* (pp. 41-53). Buckingham, UK & Philadelphia, PA: The Society for Research into Higher Education & Open University Press.

Gibbs, G., & Lucas, L. (1997). Coursework assessment, class size and student performance 1984-94. *Journal of Further and Higher Education, 21*(2), 183.

Gordijn, J., & Nijhof, W. (2002). Effects of complex feedback on computer-assisted modular instruction. *Computers and Education, 39*(2), 183-200.

Grebenik, P., & Rust, C. (2002). IT to the rescue. In P. Schwartz & G. Webb (Eds.), *Assessment. Case studies, experience and practice from higher education* (pp. 18-24). London: Kogan Page.

Greenwood, D., & Levin, M. (1998). *Introduction to action research: social research for social change*. Thousand Oaks, CA: Sage.

Gretes, J. A., & Green, M. (2000). Improving undergraduate learning with computer-assisted assessment. *Journal of Research on Computing in Education, 33*(1), 46-54.

Grundy, S. (1987). *Curriculum: Product or praxis*. Lewes: Falmer.

Hale, S. E. (1964). *Report of the committee on university teaching method*. London: HMSO.

Harpp, D. N. (1994). Special problems of big classes: some common sense solutions. *Journal of Chemical Education, 71*(8), 629-636.

Hawe, E. (2003). It's pretty difficult to fail; the reluctance of lecturers to award a fail grade. *Assessment & Evaluation in Higher Education, 28*(4), 371-382.

Henry, C., & McTaggart, R. (1996). Bottom rungs - Action research. *Changing Education, 3*(2), 6-11.

Herr, K., & Anderson, G. (2005). *The action research dissertation*. Thousand Oaks, CA: Sage.

Hodgkinson, H. L. (1957). Action research - a critique. *Journal of Educational Sociology, 31*(4), 137-153.

Hunt, E., & Pellegrino, J. W. (2002). Issues, examples, and challenges in formative assessment. In D. F. Halpern & M. D. Hakel (Eds.), *Applying the science of learning to university teaching and beyond*. San Francisco, CA: Wiley.

Hussmann, S., & Bigdeli, A. (2002). *Introduction of web-based learning and communication tools in Electrical Engineering at the University of Auckland.* Paper presented at the 13th Annual Conference, Australasian Association for Engineering Education, Canberra, Australia.

Hussmann, S., & Smaill, C. (2003). The use of web-based learning and communication tools in electrical engineering. *Australasian Journal of Engineering Education, online publication 2003-01*.

Ilgen, D., & Davis, C. (2000). Bearing bad news: reactions to negative feedback. *Applied Psychology, 49*(3), 550-565.

Innis, K. (1996). *Diary survey: how undergraduate full-time students spend their time*. Leeds: Leeds Metropolitan University.

Jacob, L. C., & Chase, C. I. (1992). *Developing and using tests effectively: a guide for faculty*. San Francisco, CA: Jossey-Bass.

Johnson, D. E., & Mihal, W. L. (1973). Performance of Blacks and Whites in computerized versus manual testing environments. *American Psychologist, 28*(8), 694-699.

Kapes, J. T., & Vansickle, T. R. (1992). Comparing paper-pencil and computer-based versions of the Harrington-O'Shea career decision-making system. *Measurement and Evaluation in Counselling and Development, 25*(1), 5-13.

Kashy, D. A., Albertelli, G., Kashy, E., & Thoennessen, M. (2001). Teaching with ALN technology: benefits and costs. *Journal of Engineering Education, 90*(4), 499-505.

Kashy, E., Thoennessen, M., Albertelli, G., & Tsai, Y. (2000). Implementing a large on-campus ALN: faculty perspective. *Journal of Asynchronous Learning Networks, 4*(3), 231-244.

Kashy, E., Thoennessen, M., Tsai, Y., Davis, N. E., & Wolfe, S. L. (1997). *Using networked tools to enhance student success rates in large classes.* Paper presented at the Frontiers in Education, Pittsburgh.

Kember, D., & Kelly, M. (1993). *Improving teaching through action research*. Campbelltown, Australia: Higher Education Research and Development Society of Australasia.

Kemmis, S., & McTaggart, R. (1988). *The action research planner*. Melbourne, Australia: Deakin University Press.

Kemmis, S., & McTaggart, R. (1992). *The actioin research planner*. Geelong: Deakin University Press.

Kluger, A., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.

Kortemeyer, G., Albertelli, G., Bauer, W., Berryman, F., Bowers, J., Hall, M., et al. (2003). *The learning online network with computer-assisted personalized approach (lon-capa).* Paper presented at the Computer based learning in science conference, Cyprus.

Kruhlak, R., Mullins, T., Coghill, C., & Ng, H. K. (2005a). *Online Assessment in First Year Physics.* Paper presented at the Uniserve Science First Year Experience Forum: Assessment, Sydney, Australia.

Kruhlak, R., Mullins, T., Coghill, C., & Ng, H. K. (2005b). *Online Practice and Assessment in First Year Physics.* Paper presented at the Canadian Association of Physicists Congress, Vancouver, Canada.

Kuhn, T. (1970). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research, 47*(1), 211-232.

Kulik, J. A., & Kulik, C.-L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*(1), 79-97.

Lambert, D., & Lines, D. (2000). *Understanding assessment*. London: RoutledgeFalmer.

Lather, P. (1986). Research as praxis. *Harvard Educational Review, 56*(3), 257-277.

Lee, G., & Weerakoon, P. (2001). The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based and paper-and-pen multiple-choice tests. *Medical Teacher, 23*(2), 152-157.

Lee, J. A., & Hopkins, L. (1985). *The effects of training on computerized aptitude test performance and anxiety.* Paper presented at the 56th annual meeting of the Eastern Psychological Association, Boston, MA.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues, 2*(4), 34-46.

Lewin, K. (1948). *Resolving social conflicts*. New York: Harper.

Lewis-Beck, M. (Ed.). (1993). *Regression analysis* (Vol. 2). London: Sage.

Lian, Y., & Hu, Y. (2005). Enabling learning-by-doing in a large class with the help of an e-learning platform. In W. Aung, R. W. King, J. Moscinski, S. Ou & L. M. S. Ruiz (Eds.), *Innovations 2005. World innovations in engineering education and research*. New York: Begell House.

Light, G., & Cox, R. (2001). *Learning and teaching in higher education*. London: Paul Chapman.

Linn, R., & Gronlund, N. (2000). *Measurement and assessment in teaching*. Upper Saddle River, NJ: Merrill.

Llabre, M. M., Clements, N. E., Fitzhugh, K. B., & Lancelotta, G. (1987). The effect of computer-administered testing on test anxiety and performance. *Journal of Educational Computing Research, 3*(4), 429-433.

LON-CAPA. (2006). From http://www.lon-capa.org/index.html

Mafi, M. (1989). Involving students in a time-saving solution to the homework problem. *Engineering Education, 79*(3), 444-446.

Mason, R. (1998). *Globalising education: Trends and applications*. London: Routledge Studies in Distance Education.

Mayes, D. K., Sims, V. K., & Koonce, J. M. (2001). Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics, 28*(6), 367-378.

McDonald, A. (2002). The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessments. *Computers and Education, 39*(3), 299-312.

McDowell, L. (1995). Effective teaching and learning on foundation and access courses in engineering, science and technology. *European Journal of Engineering Education, 20*(4), 417-425.

McMillan, J. H. (2004). *Classroom assessment*. Boston: Pearson Education.

McNiff, J., Lomax, P., & Whitehead, J. (2003). *You and your action research project*. London: RoutledgeFalmer.

McTaggart, R. (1991). *Action research: A short modern history*. Geelong: Deakin University.

Mead, A., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: a meta-analysis. *Psychological Bulletin, 114*(3), 449-485.

Mehta, S. I., & Schlecht, N. W. (1998). Computerized assessment technique for large classes. *Journal of Engineering Education, 87*(2), 167-172.

Merceron, A., & Yacef, K. (2003). *A Web-based tutoring tool with mining facilities to improve learning and teaching.* Paper presented at the 11th International Conference on Artificial Intelligence in Education, Sydney, Australia.

Miller, C. M. L., & Parlett, M. (1974). *Up to the mark: A study of the examination game*. Guildford, UK: Society for Research into Higher Education.

Minaei-Bidgoli, B. (2005). Data mining for a web-based educational system. from http://www.lon-capa.org/papers/BehrouzThesisRevised.pdf

Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (2003). *Predicting student performance: an application of data mining methods with the educational web-based system LON-CAPA.* Paper presented at the Frontiers in education, Boulder, CO.

N6. (2006). From http://www.qsrinternational.com/products/productoverview/product_overview.htm

Nafalski, A., & Samaan, N. (1997). *Computerised assessment in engineering disciplines.* Paper presented at the Australasian Association for Engineering Education 9th Annual Conference, Ballarat, Australia.

nDeva. (2006). From https://ndeva.auckland.ac.nz/ndeva/

Neuman, G., & Baydoun, R. (1998). Computerization of paper-and-pencil tests: when are they equivalent? *Applied Psychological Measurement, 22*(1), 71-83.

Ng, H. K. (2004). OASIS performance investigation: Unpublished part-four project report, University of Auckland, New Zealand.

Noll, V., Scannell, D., & Craig, R. (1979). *Introduction to educational measurement*. Boston, MA: Houghton-Mifflin.

Nordhaus, W. (2002). *Progress in computing.* Paper presented at the Workshop on Economic Measurement: Hedonic Price Indexes: Too Fast? Too Slow? Or Just Right?, Washington, DC.

O'Reilly, U., Alexander, S., Sweeney, P., & McAllister, G. (2002). Utilizing automated assessment for large student cohorts. In W. Aung, P. Hicks, L. Scavarda, V. Roubicek & C.-H. Wei (Eds.), *Engineering education and research - 2001: A chronicle of worldwide innovations*. Arlington, VA,: iNEER in cooperation with Begell House.

Palmer, S., & Bray, S. L. (2001). Longitudinal study of computer usage in flexible engineering education. *Australian Journal of Educational Technology, 17*(3), 313-329.

Pascarella, A. M. (2002). *CAPA (Computer-assisted personalized assignments) in a large university setting.* Unpublished PhD, University of Colorado, Boulder.

Paxton, M. (2000). A linguistic perspective on multiple choice questioning. *Assessment & Evaluation in Higher Education, 25*(2), 109-119.

Peat, M., & Franklin, S. (2003). Has student learning been improved by the use of online and offline formative assessment opportunities? *Australian Journal of Educational Technology, 19*(1), 87-99.

Questionmark. (2006). Questionmark perception for web. From www.questionmark.com

Raikes, N., & Harding, R. (2003). The horseless carriage stage: replacing conventional measures. *Assessment in Education, 10*(3), 267-277.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science, 28*, 4-13.

Reinhardt, A. (1995, 1995, March). New ways to learn. *Byte, The Magazine of Technology Integration, 20,* 50-72.

Ricketts, C., & Wilks, S. J. (2002). Improving student performance through computer-based assessment: insights from recent research. *Assessment & Evaluation in Higher Education, 27*(5), 475-479.

Robbins, L. (1963). *Higher education*. London: HMSO.

Robson, C. (2002). *Real world research*. Oxford: Blackwell.

Rocklin, T., & Thompson, J. (1985). Interactive effects of test anxiety, test difficulty, and feedback. *Journal of Educational Psychology, 77*(3), 368-372.

Rogers, E. M. (1969). Examinations: Powerful agents for good or ill in teaching. *American Journal of Physics, 37*, 954 - 962.

Rose, R. (2002). Teaching as a "research-based profession": encouraging practitioner research in special education. *British Journal of Special Education, 29*(1), 44-48.

Rothberg, S., Lamb, F., & Wallace, A. (2001). Computer assisted learning in engineering degree programmes: a survey at the end of the 20th century. *International Journal of Engineering Education, 17*(6), 502-511.

Russell, M. (1999). Testing on computers: a follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives, 7*(20).

Russell, M., Goldberg, A., & O'Connor, K. (2003). Computer-based testing and validity: a look back into the future. *Assessment in Education, 10*(3), 279-293.

Russell, M., & Haney, W. (1997). Testing writing on computers: an experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives, 5*(3).

Russell, M., & Plati, T. (2002). Does it matter with what I write? comparing performance on paper, computer and portable writing devices. *Current Issues in Education, 5*(4).

Sabar, N. (2002). Towards principled practice in evaluation: learning from instructors' dilemmas in evaluating graduate students. *Studies in Educational Evaluation, 28*(4), 329-345.

Sadler, D. (1983). Evaluation and the improvemnt of academic learning. *Journal of Higher Education, 54*, 60-79.

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144.

Sambell, K., & McDowell, L. (1998). The construction of the hidden curriculum: messages and meanings in the assessment of student learning. *Assessment & Evaluation in Higher Education, 23*(4), 391-402.

Sambell, K., Sambell, A., & Sexton, G. (1999). Student perceptions of the learning benefits of computer-assisted assessment: a case study in electronic engineering. In S. Brown, P. Race & J. Bull (Eds.), *Computer-assisted assessment in higher education*. London: Kogan Page.

Sapp, M. (1993). *Test anxiety: Applied research, assessment, and treatment intervention*. Lanham, MD.: University Press of America.

Schimmel, B. J. (1983). *A meta-analysis of feedback to learners in computerized and programmed instruction.* Paper presented at the American Educational Research Association, Montreal.

Schon, D. A. (1991). *The reflective practitioner: How professionals think in action*. Aldershot: Avebury.

Schon, D. A. (1995). Knowing-in-action: The new scholarship requires an new epistemology. *Change, 27*(6), 27-34.

Scott, N. W., & Stone, B. J. (1998). A flexible web-based tutorial system for engineering, maths and science subjects. *Global Journal of Engineering Education, 2*(1), 7-16.

Scott, N. W., & Stone, B. J. (1999). We did it our way (and you must do it your way). *Australasian Journal of Engineering Education, 8*(2), 99-123.

Scriven, M. (1967). *The methodology of evaluation* (Vol. No. 1 pp. 39-83). Chicago, IL: Rand McNally.

Seale, J., Chapman, J., & Davey, C. (2000). The influence of assessments on students' motivation to learn in a therapy degree course. *Medical Education, 34*(8), 614-621.

Shermis, M., & Lombard, D. (1998). Effects of computer-based test administrations on test anxiety and performance. *Computers in Human Behavior, 14*(1), 111-123.

Siege. (2006). From http://www.joedog.org/

Sim, G., Holifield, P., & Brown, M. (2004). Implementation of computer assisted assessment: lessons from the literature. *ALT-J, Research in Learning Technology, 12*(3), 215-229.

Slavin, R. (1978). Separating incentives, feedback and evaluation: toward a more effective classroom system. *Educational Psychologist, 13*, 97-100.

Sly, L. (1999). Practice tests as formative assessment improve student performance on computer-managed learning assessments. *Assessment & Evaluation in Higher Education, 24*(3), 339-343.

Sly, L. (2000). *Computer managed learning assessment in higher education: the effect of a practice test.*, Curtin University of Technology, Perth, Australia.

Sly, L., & Rennie, L. (1999). Computer managed learning as an aid to formative assessment in higher education. In S. Brown, P. Race & J. Bull (Eds.), *Computer-assisted assessment in higher education* (pp. 113-120). London: Kogan Page.

Smaill, C. (2004). *The implementation and evaluation of a Web-based learning and assessment tool in electrical engineering.* Paper presented at the association for engineering education in southeast and east Asia and the Pacific (AEESEAP) mid-term conference, Auckland.

Smaill, C. (2005). The implementation and evaluation of OASIS: a Web-based learning and assessment tool for large classes. *IEEE Transactions on Education, 48*(4), 658-663.

Smaill, C., & Hussmann, S. (2003). *The implementation and evaluation of an individualised, web-based, formative and summative assessment software tool for large classes.* Paper presented at the Fourth International Conference on Information Technology Based Higher Education and Training, Marrakech, Morocco.

Snyder, B. R. (1971). *The hidden curriculum*. New York: Knopf.

Soderstand, M. A. (1994). The new electrical and computer engineering curriculum at University of California-Davis. *IEEE Transactions on Education, 37*(2), 136-146.

Sporn, B. (1999). *Adaptive university structures: An analysis of adaptation to socioeconomic environments of US and European universities.* London: Jessica Kingsley.

SPSS. (2006). From http://www.spss.com/

Stephens, D., Bull, J., & Wade, W. (1998). Computer-assisted assessment: suggested guidelines for an institutional strategy. *Assessment & Evaluation in Higher Education, 23*(3), 283-294.

Susman, G. I., & Evered, R. D. (1978). An assessment of the scientific merits of action research. *Administrative Science Quarterly, 23*(4), 582-603.

Tandon, S. D., Kelly, J. G., & Mock, L. (2001). Participatory action research as a resource for developing African American community leadership. In D. Tolman & M. Brydon-Miller (Eds.), *From subjects to subjectivities: A handbook of interpretive and participatory methods* (pp. 200-217). New York: New York University Press.

Tartaglia, A., & Tresso, E. (2002). An automatic evaluation system for technical education at the university level. *IEEE Transactions on Education, 45*(3), 268 - 275.

Thoennessen, M., & Harrison, M. J. (1996). Computer-assisted assignments in a large physics class. *Computers and Education, 27*(2), 141-147.

Thoennessen, M., Kashy, E., Tsai, Y., & Davis, N. E. (1997). *Application of technology and asynchonous learning networks in large lecture classes.* Paper presented at the International conference on system sciences, Hawaii.

Tseng, H., Tiplady, B., Macleod, H., & Wright, P. (1998). Computer anxiety: a comparison of pen-based personal digital assistants, conventional computer and paper assessment of mood and performance. *British Journal of Psychology, 89*(4), 599-610.

Vogel, L. (1994). Explaining performance on P&P versus computer mode of administration for the verbal section of the Graduate Record Exam. *Educational Computing Research, 11*(4), 369-383.

Watkins, K. (1991). *Validity in action research.* Paper presented at the Annual meeting of the American Educational Research Association, Chicago.

WebCT. (2006). From http://www.webct.com

Weinberg, S., & Abramowitz, S. (2002). *Data analysis for the behavioral sciences using SPSS*. Cambridge: Cambridge University Press.

Whyte, W. F., Greenwood, D. J., & Lazes, P. (1991). Participatory action research: Through practice to science in social research. In W. F. Whyte (Ed.), *Participatory action research*. Newbury park, CA: Sage.

Wise, S. L., & Plake, B. S. (1989). Research on the efects of administering tests via computers. *Educational Measurement: Issues and Practice, 8*(3), 5-10.

Wolf, D. P. (1993). Assessment as an episode of learning. In R. E. Bennett & W. C. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 213-240). Hillsdale, NJ: Erlbaum.

Wong, E. D. (1995). Challenges confronting the researcher/teacher: conflicts of purpose and conduct. *Educational Researcher, 24*(3), 22-28.

Wragg, E. C. (1994). *An introduction to classroom observation*. London: Routledge.

Zakrzewski, S. (1999). A structured approach to mass implementation of computer-based assessment (CBA). In S. Brown, P. Race & J. Bull (Eds.), *Computer-assisted assessment in higher education*. London: Kogan Page.

Zakrzewski, S., & Bull, J. (1998). The mass implementation and evaluation of computer-based assessments. *Assessment & Evaluation in Higher Education, 23*(2), 141-152.

Zakrzewski, S., & Steven, C. (2000). A model for computer-based assessment: the catherine wheel principle. *Assessment & Evaluation in Higher Education, 25*(2), 201-215.

Zuber-Skerritt, O. (1996a). Emancipatory action research for organisational change and management development. In O. Zuber-Skerritt (Ed.), *New directions in action research* (pp. 83-105). London: Falmer Press.

Zuber-Skerritt, O. (1996b). Introduction. In O. Zuber-Skerritt (Ed.), *New directions in action research* (pp. 3-9). London: Falmer Press.

Appendix one: OASIS survey form

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

FFQ-Formative Feedback Questionnaire-Course

Circuits and Systems

ELECTENG 202 S1 C

Survey#:5216

Please shade in the bubble that is closest to your response along a scale from Strongly Disagree ... Strongly Agree. If you feel that the item is not relevant to your experience of this course or teacher shade in Not Applicable. Your responses will be used by the lecturer or department to reflect on the quality of teaching or courses, and make improvements if necessary. Your participation in this survey is voluntary and anonymous. You are entitled to see the results of this survey.

Fill in the bubble completely with black/ blue pen or a pencil.

Course group or section (if applicable)

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Not applicable
1. OASIS is easy to use	○	○	○	○	○	○
2. I often use OASIS	○	○	○	○	○	○
3. OASIS helped improve my skill level	○	○	○	○	○	○
4. I am more confident about my learning after using OASIS	○	○	○	○	○	○
5. The problems provided in OASIS helped me understand the course material better	○	○	○	○	○	○
6. OASIS helps me to prepare for the assessments	○	○	○	○	○	○
7. I come to campus less often because of OASIS	○	○	○	○	○	○
8. I like the instant performance feedback using OASIS	○	○	○	○	○	○
9. I prefer to do problems from the text rather than OASIS problems	○	○	○	○	○	○
10. It would be a good idea to have OASIS in other courses	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○
	○	○	○	○	○	○

Office use only

PLEASE TURN OVER FOR OPEN-ENDED ITEMS

29057

What do you like most about OASIS?
What do you like least about OASIS?
What changes or improvements would you suggest for OASIS?
What benefits do you think students get from using OASIS?
Have you any other comments to offer about OASIS?
Thank you for your comments

Appendix two: information sheets and permission forms

Information sheet for all students

OASIS: Online Assessment System with Integrated Study

OASIS is a web-based tutorial and assessment software package currently being used in the Department of Electrical & Computer Engineering. You can access OASIS from a university computer or anywhere you have the internet. When you practise problems on OASIS, you will receive prompt feedback as well as the opportunity to repeat problems, with different numerical values each time, until you are satisfied you grasp the solution method. OASIS can also deliver and mark assignments and tests electronically.

Evidence suggests that students benefit from using OASIS. With the approval of the HOD, I am carrying out some course evaluation research to find out more about this. I will be studying how you use OASIS and how it helps you learn. This knowledge will help me modify OASIS and the way it is used to further improve student learning. OASIS will record the problems you try and your success rate on them. This will help you keep track of which problems you have mastered and which problems need further attention. My research will involve a statistical analysis of your OASIS data, and test and examination results. Be reassured that this data will only be used in a statistical way in my research. No one will be able to identify you or any other individuals from my research.

Later in the course, I will ask you to complete one or two anonymous questionnaires about OASIS as part of your course evaluation. Your participation in this is voluntary. What I find out may help me improve the way this and other courses are organised.

Thank you very much for being part of this study. We should all benefit from it. Please contact me if you have any queries or wish to know more.

Chris Smaill Tel. 3737599 extn 83012 Email c.smaill@auckland.ac.nz

If you have any concerns of an ethical nature you can contact the Chair of the University of Auckland Human Participants Ethics Committee at 3737599 extn 87830.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE on 11.5.2005 for a period of three years, from 26.06.2005 to 26.06.2008.
Reference 2002 / 158.

Information sheet for potential interviewees

OASIS: Online Assessment System with Integrated Study

At the start of this course you and your fellow students received an information sheet about OASIS. Since then there has been considerable student use of OASIS. As part of my research into the use of OASIS I have randomly selected some people to interview and your name came up. It is your choice whether you are interviewed. Either way, your grades and your academic standing will not be affected.

The purpose of the interview is to find out how you feel about the use of OASIS as part of your learning. Interviews would take about half an hour. What you say may help me improve the way this and other courses are organised. I would like, with your consent, to audiotape interviews. The tape could be turned off at any time. You will be able to withdraw information, without stating a reason, at any time up to six months after the interview. Please be reassured that no one will be able to identify any individuals from my research. All information you provide in an interview is confidential and your name will not be used.

If you do wish to be interviewed please let me know by filling in the Consent Form and returning it to me. Thank you very much for your time and help in making this study possible. We should all benefit from it. Please contact me if you have any queries or wish to know more.

Chris Smaill
Tel. 3737599
extn 83012
Email c.smaill@auckland.ac.nz

My Head of Department is: Professor Allan Williamson
Tel. 37377599
extn 87922

My research is for a Doctorate through Curtin University, Perth, Australia. My supervisor at Curtin is Professor David Treagust. The research programme title is: “The use of an Online Assessment System with Integrated Study (OASIS) and other approaches to enhance the learning of electrical engineering”.

If you have any concerns of an ethical nature you can contact the Chair of the University of Auckland Human Participants Ethics Committee at 3737599 extn 87830.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE on 11.5.2005 for a period of three years, from 26.06.2005 to 26.06.2008.
Reference 2002 / 158.

Information sheet for Head of Department

OASIS: Online Assessment System with Integrated Study

OASIS is a web-based tutorial and assessment software package currently being used in the Department of Electrical & Computer Engineering. Students can access OASIS from a university computer or anywhere they have internet access. When they practise problems on OASIS, they receive prompt feedback as well as the opportunity to repeat problems, with different numerical values each time, until they are satisfied they grasp the solution method. OASIS can also deliver and mark assignments and tests electronically. This latter is of considerable benefit to staff.

Evidence suggests that students benefit from using OASIS. I am carrying out some research to find out more about this. My research is for a Doctorate through Curtin University, Perth, Australia. My supervisor at Curtin is Professor David Treagust. The research programme title is: "The use of an Online Assessment System with Integrated Study (OASIS) and other approaches to enhance the learning of electrical engineering".

I will be studying how students use OASIS and how it helps them learn. This knowledge will help me modify OASIS and the way it is used to further improve student learning. OASIS will record the problems students try and their success rates. My research will involve a statistical analysis of student OASIS data, and test and examination results. This data will only be used in a statistical way in my research. No one will be able to identify any individuals from my research. The data used will not go beyond the data available to any lecturer who uses OASIS with his or her class.

Students will be asked, on a voluntary basis, to complete one or two anonymous questionnaires about OASIS. What I find may help me improve the way courses are organised. Again, the questions will not go beyond the sort of question routinely asked in course evaluations.

I will also randomly select some students to approach for interview. The purpose of the interviews is to find out how students feel about the use of OASIS as part of their learning. Of course, the consent of the student would be obtained prior to interview, and also for audiotaping. All information provided in an interview would be confidential and student names would not be used.

I have obtained University of Auckland Ethics Committee approval for this research.

Chris Smaill

If you have any concerns of an ethical nature you can contact the Chair of the University of Auckland Human Participants Ethics Committee at 3737599 extn 87830.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE on 11.5.2005 for a period of three years, from 26.06.2005 to 26.06.2008.
Reference 2002 / 158.

OASIS CONSENT FORM, INTERVIEWEE

This consent form will be held securely for a period of six years.

Research Title: The use of an Online Assessment System with Integrated Study (OASIS) and other approaches to enhance the learning of electrical engineering

Researcher: Chris Smaill

I have been given and have understood an explanation of this research project. I have had an opportunity to ask questions and have them answered.

I understand that I may withdraw myself or any information traceable to me at any time up to six months after the date below without giving a reason.

- I agree to take part in this research.
- I agree/do not agree that the interview will be audio taped.

Signed:

Name:
(please print clearly)

Email:

Date:

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE on 11.5.2005 for a period of three years, from 26.06.2005 to 26.06.2008.
Reference 2002 / 158.

OASIS CONSENT FORM, HOD

Research Title: The use of an Online Assessment System with Integrated Study (OASIS) and other approaches to enhance the learning of electrical engineering

Researcher: Chris Smaill

I have discussed this research project with the researcher and I have read some of its documentation, in particular the “Information sheet for Head of Department”.

I give my permission for the researcher to carry out a statistical analysis of student OASIS data, and test and examination results. I understand that this data will only be used in a statistical way and that no one will be able to identify any individuals from the research.

Signed:

Professor Allan Williamson

Head of Department

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE on 11.5.2005 for a period of three years, from 26.06.2005 to 26.06.2008.
Reference 2002 / 158.

Appendix three: OASIS practice; screenshots of representative student pages

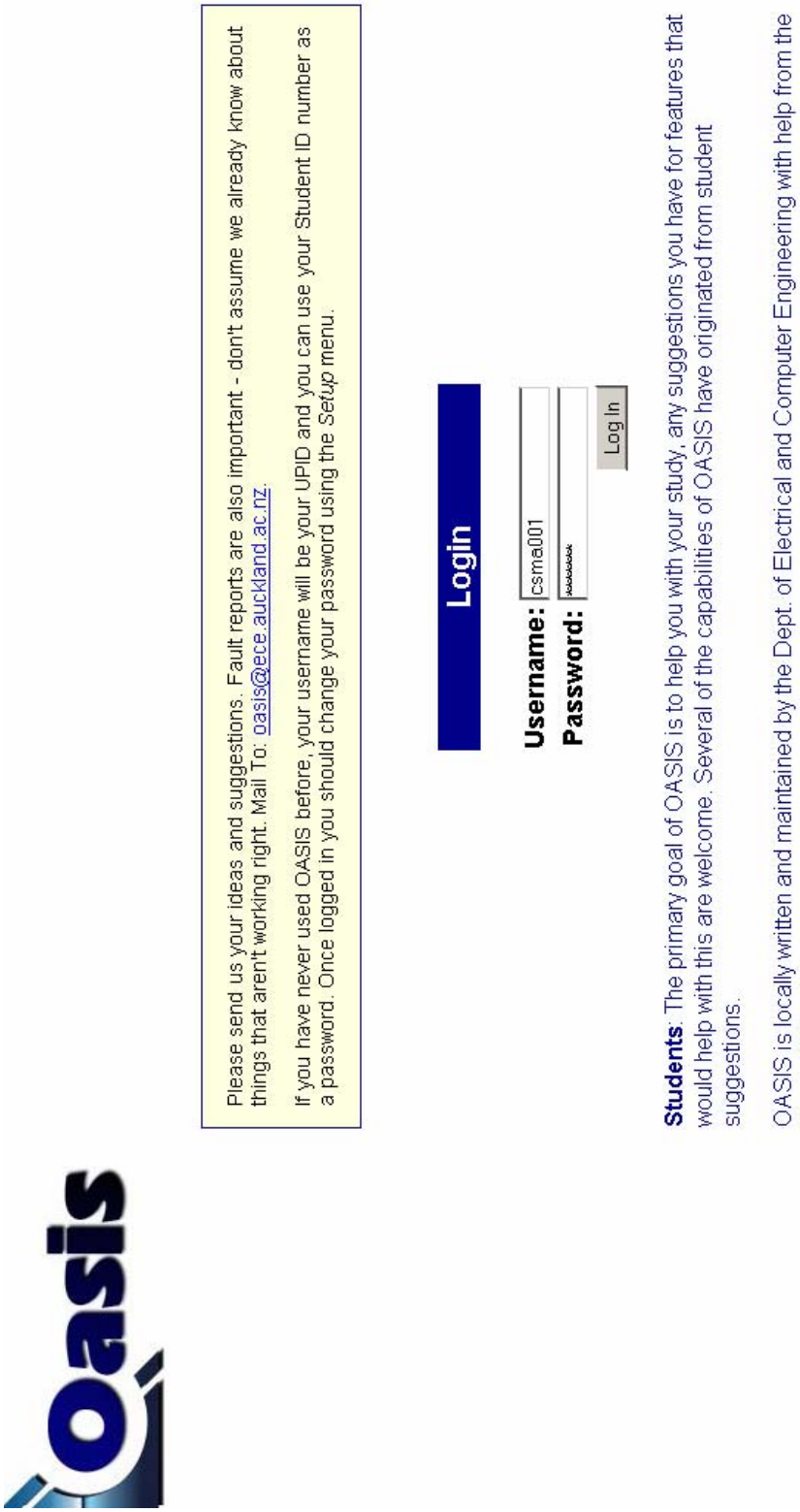


Figure A-3-1: OASIS login page for both students and staff

Home News Practice Assess Setup Logout

Status for csma001

You are a registered user.

Welcome to Oasis

(version 3.6.5)

Apologies for problems in the last couple of days, the OASIS server suffered a hardware failure and had to be replaced at short notice. Things should be stable now.

Please report any problems or suggestions you have to oasis@ece.auckland.ac.nz.

Oasis contains several sections:

News

The latest news and information about Oasis.

Practice

Allows you to practice solving a near endless supply of exercises. Oasis will mark these for you so you know how you are doing.

Assess

This is where you should go for terms-tests and assignments. Work under this area may contribute to your real coursework marks, so be sure to **practice** first.

Setup

Here you can change your password. Teaching staff can also set up new questions, administer assignments, and perform related tasks.

You may also **Logout** to leave Oasis.

Figure A-3-2: OASIS main menu for both students and staff

Home News Practice Assess Setup Logout

Practice **Choose A Course**

Course	Title
COMPSYS202	Object Oriented Design and Programming
ELECTENG101	Electrical Engineering Systems
ELECTENG202	Circuits and Systems
ELECTENG207	Microelectronic Circuits
ELECTENG208	Electric Circuit Analysis
ELECTENG302	Engineering Electromagnetics II
ELECTENG303	Systems and Control
ELECTENG305	Electronic Devices and Technology
ELECTENG307	Transmission Lines and Systems
EXAMPLE101	Some Examples
MECHENG270	Software Design
OASIS Demo	A set of demonstration questions.
PHYSICS120	Physics of Energy
PHYSICS150	Physics of Technology
PHYSICS160	Physics for the Life Sciences
SOFTENG206	Software Development Craft

Figure A-3-3: OASIS practice menu showing courses that offer OASIS practice

Figure A-3-4: OASIS sub-category practice menu showing topics available for OASIS practice within the course 'Circuits & Systems', ELECTENG 202. The number of questions available for student practice in each topic is also displayed.

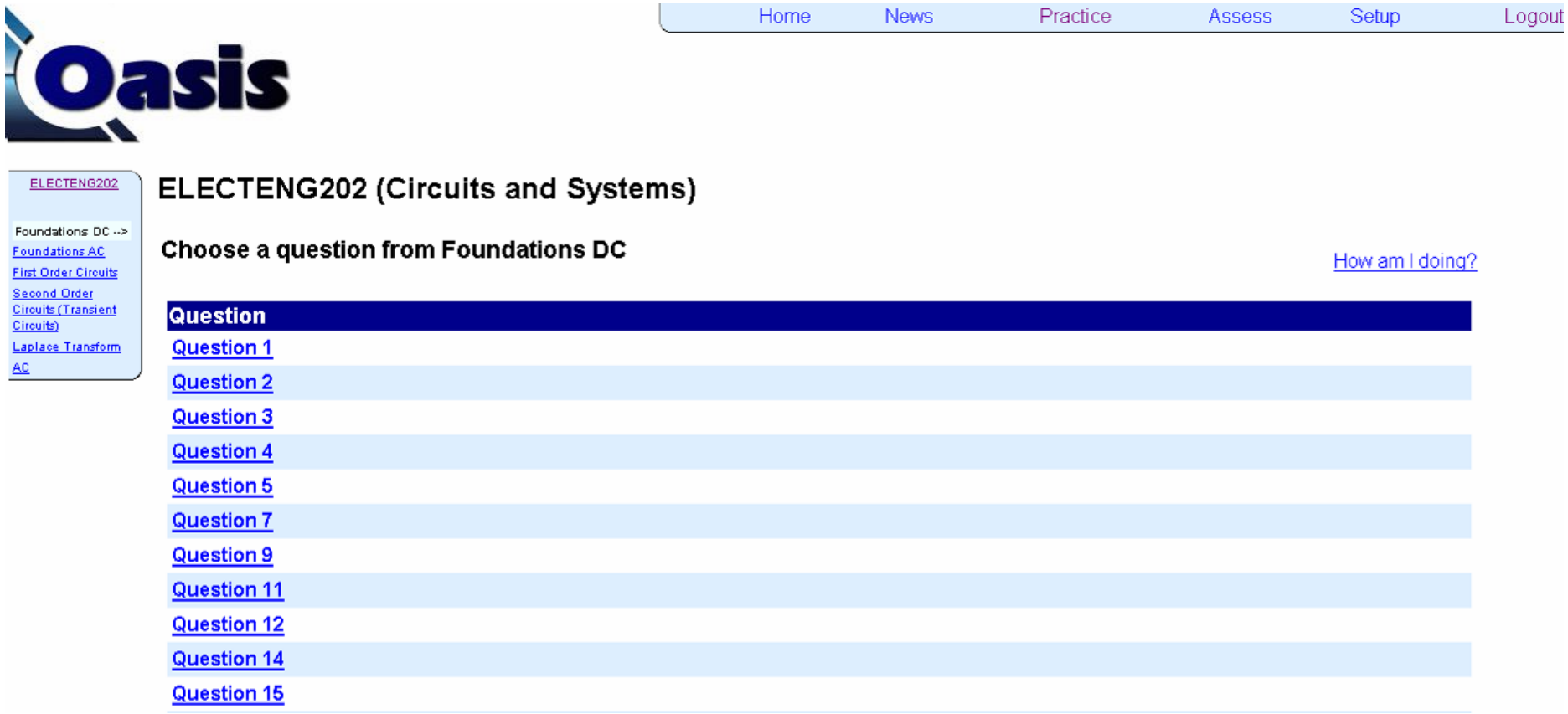


Figure A-3-5: OASIS question menu showing some of the 37 questions available for practice in the 'Foundations DC' topic in the course 'Circuits & Systems'. Some questions, such as Question 8, are hidden: they are reserved for assessment.

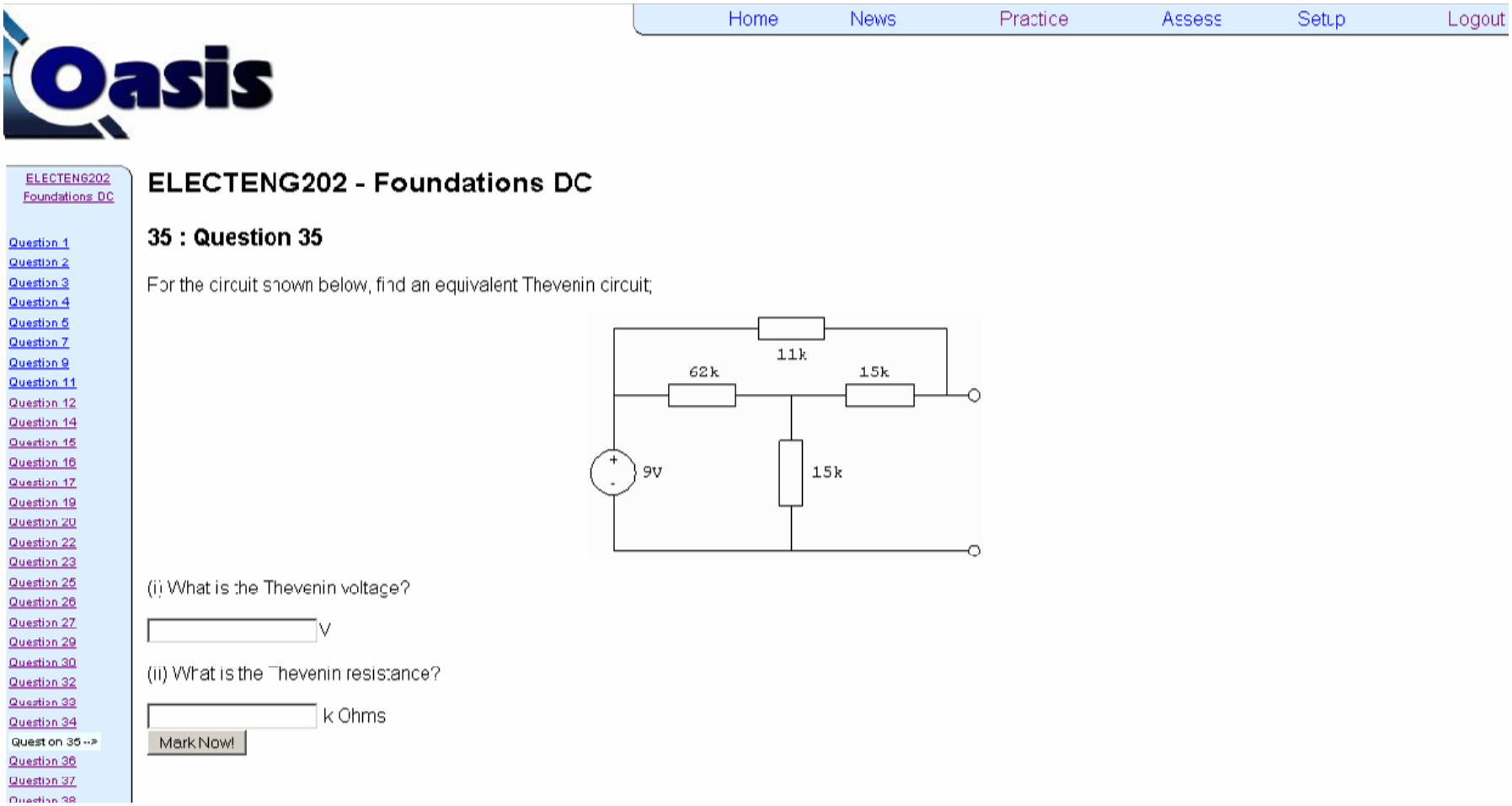


Figure A-3-6: One of the practice questions available in the 'Foundations DC' topic.

Home News Practice Assess Setup Logout

ELECTENG202
Foundations DC

Question 1
Question 2
Question 3
Question 4
Question 5
Question 7
Question 9
Question 11
Question 12
Question 14
Question 15
Question 16
Question 17
Question 19
Question 20
Question 22
Question 23
Question 25
Question 26
Question 27
Question 29
Question 30
Question 32

ELECTENG202 - Foundations DC

Previous Try Again Next

35 : Question 35

Part	Your Answer	Correct Answer	Tolerance	Marks	Comment
1	6.9	6.91	1%	1.0	Correct
2	8.3	7.82	1%	0.0	Incorrect
				Total:	1.0

Figure A-3-7: Result of answer submission to the question shown in Figure A-3 -6.

Appendix four: OASIS assessment; screenshots of representative student pages

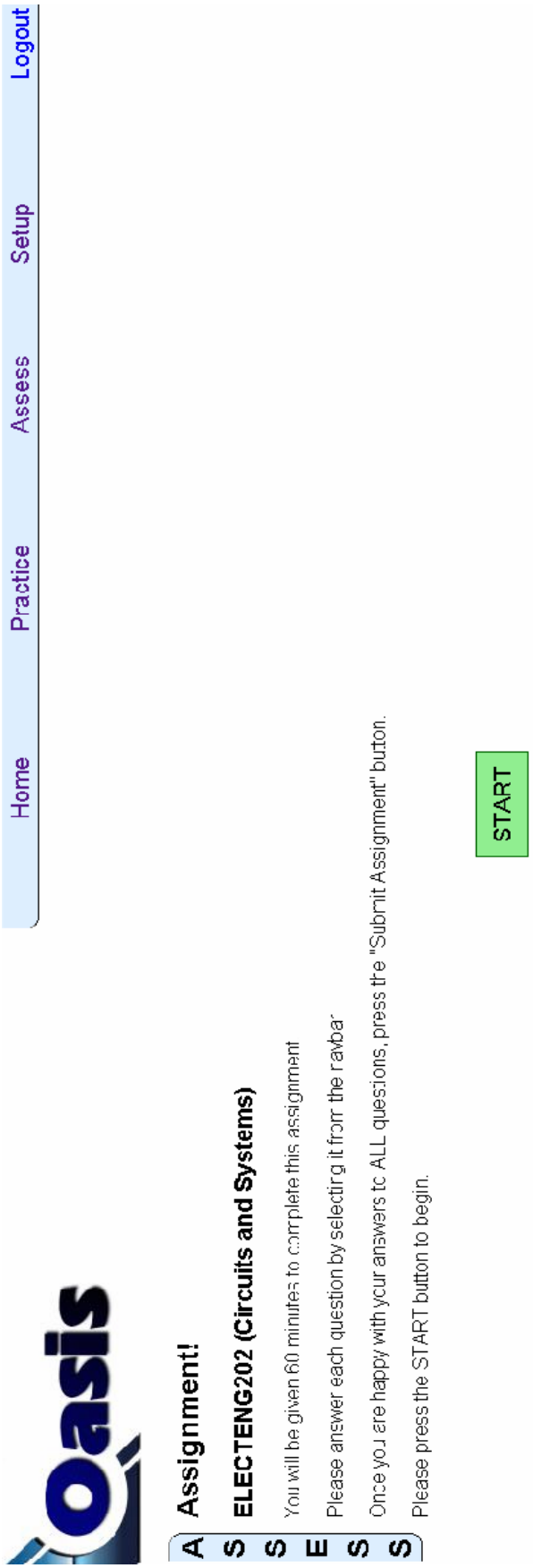


Figure A-4-1: Start page for assessments. In this case the assessment is an assignment in 'Circuits & Systems'.

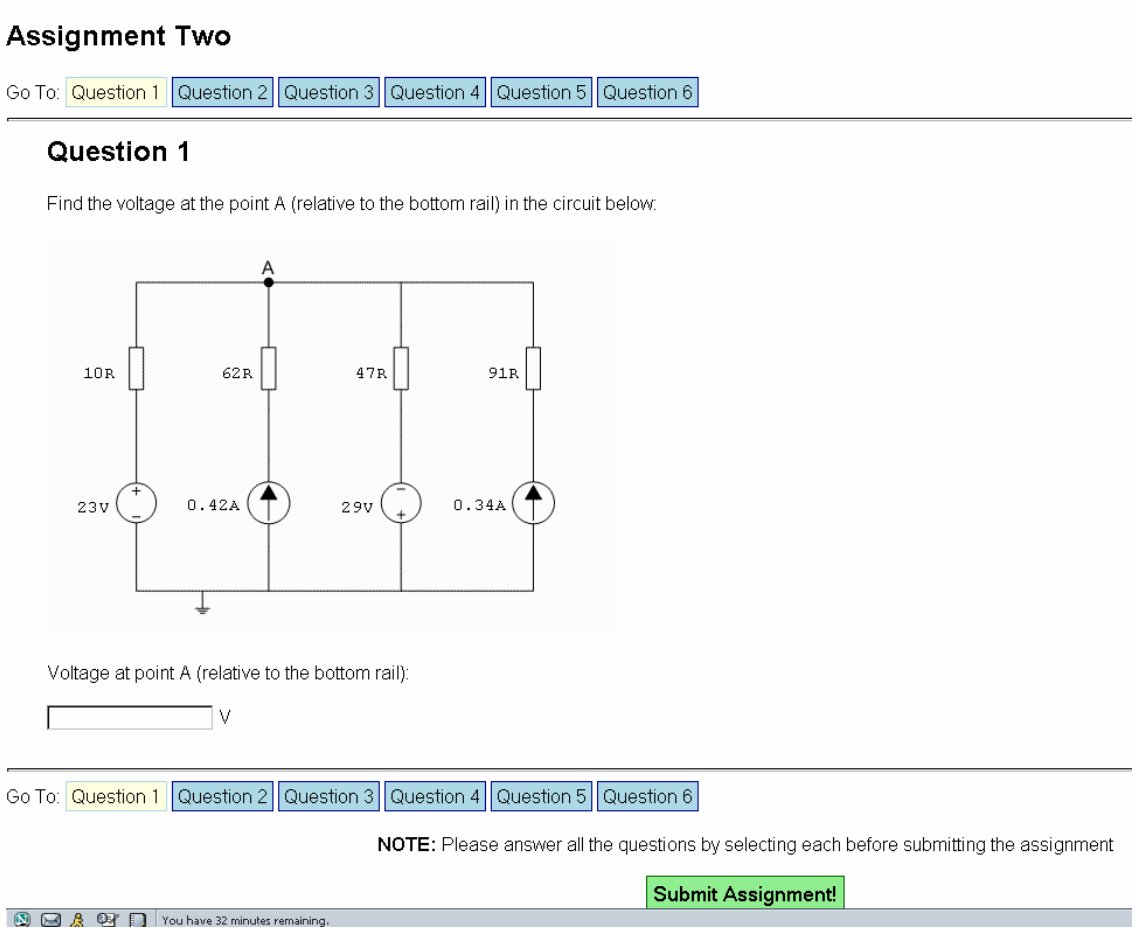


Figure A-4-2: The assignment displayed to a student after he or she clicks the 'start' button in Figure A-4-1. Note that the time remaining is displayed at the bottom of the screen.

Appendix five: OASIS staff-only pages, representative screenshots

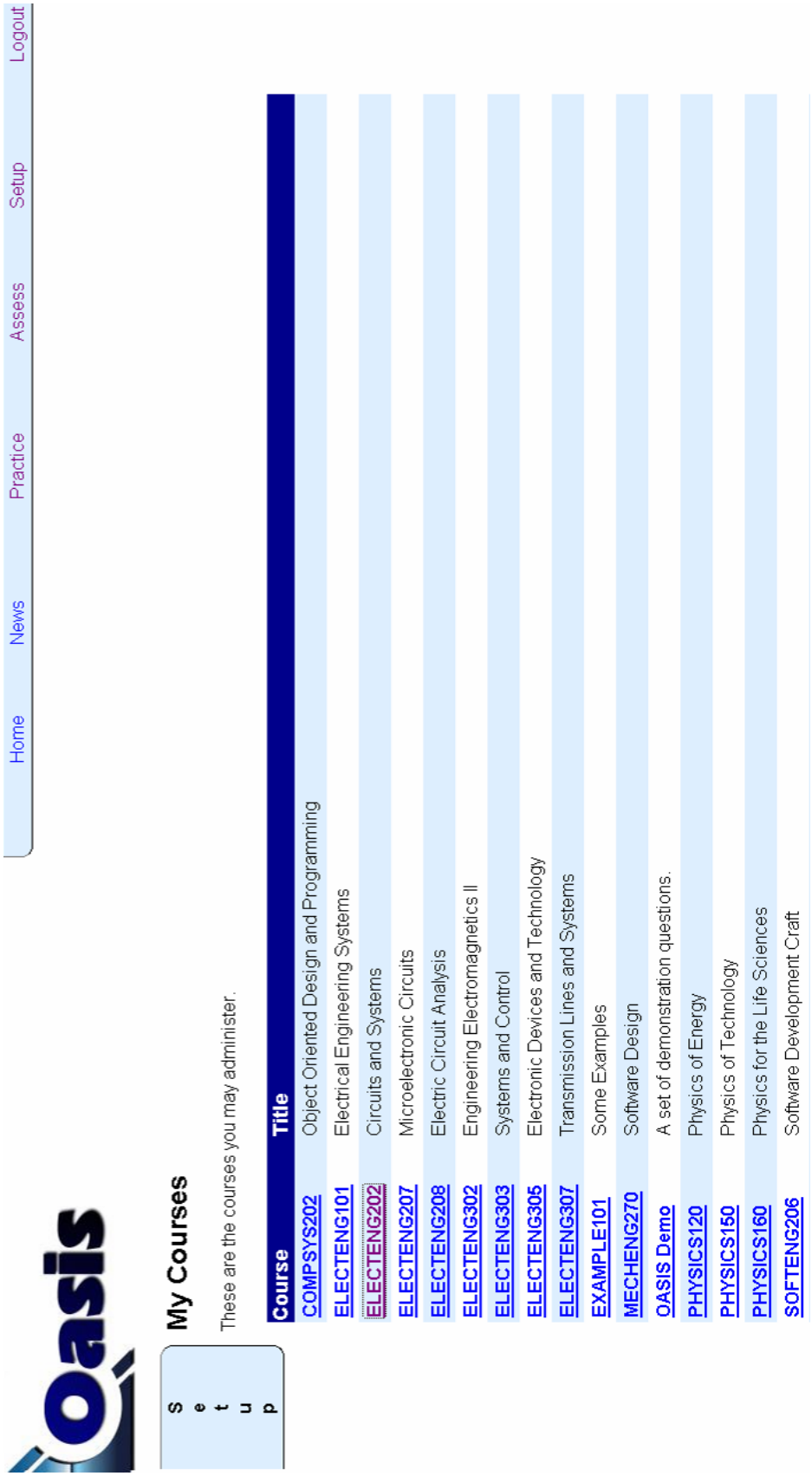


Figure A-5-1: The main setup menu. This menu is different for different staff, depending on their access rights to courses.

Home News Practice Assess Setup Logout

Oasis

Setup

Question Database

ELECTENG208 (ELECTENG208)

This class has practiced 7060 questions overall (3979 in 2006).

Category Title	Visibility	Practiced (2006)	Edit
Foundations DC	Student	4499 (2838)	Questions
Foundations AC	Student	2061 (1131)	Questions
First Order Circuits	Staff (including TAs)	278 (0)	Questions
Second Order Circuits (DC Analysis)	Staff (including TAs)	116 (0)	Questions
Laplace Transform	Staff (including TAs)	96 (0)	Questions
Op Amps	Staff (including TAs)	10 (10)	Questions
[New Category]		Save Changes	

Assessments

09 Mar	**Assignment One (View Marks)**	Overall Practice Data	Individual Practice Data	Edit
21 Mar	**Assignment Two (View Marks)**	Overall Practice Data	Individual Practice Data	Edit

Create New Assessment

Figure A-5-2: Part of the setup menu for a particular course, ELECTENG 208. Available by scrolling down, but not shown here, are links to statistical information and student search facilities.

Home News Practice Assess Setup Logou

Oasis

Setup

Statistics

on Live Questions Practice Data

ELECTENG101

Introduction to Electric Circuits

Question (Position)	Attempts	Num. Students	Attempts (CLASS)	Num. Students (CLASS)	Class Mean	Std Dev	Max	Min	Average Attempt Time
1	6810	2116	1375	476	0.700	0.459	1.0	0.0	
2	5767	1991	1190	457	0.752	0.432	1.0	0.0	
3	5487	1977	1120	465	0.799	0.401	1.0	0.0	
4	4831	1907	910	434	2.796	1.665	4.0	0.0	
5	7360	1916	1404	431	0.974	0.914	2.0	0.0	
6	5352	1885	1115	438	1.647	1.349	3.0	0.0	
7	5170	1845	1110	432	1.850	1.311	3.0	0.0	
8	3854	1734	748	394	1.533	0.818	2.0	0.0	
9	3505	1687	670	373	2.290	1.198	3.0	0.0	
10	3909	1716	758	388	1.421	0.870	2.0	0.0	
11	4829	1822	991	437	1.934	1.349	3.0	0.0	
12	7480	1915	1238	382	0.502	0.741	2.0	0.0	

Figure A-5-3: Screenshot showing information on student practice submissions for a particular course, ELECTENG 101. Information denoted by CLASS refers to current year, other data to all years combined.

Home News Practice Assess Setup Logou

Oasis

Setup

Return to Course Page

Assignment One

ELECTENG208 (36/38 students)

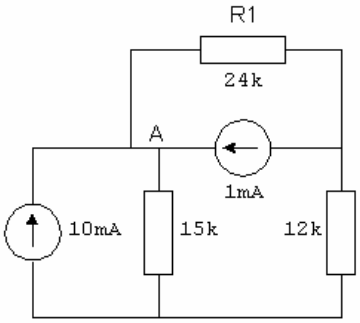
ID	Full Name	netid	Q1	Q2	Q3	Q4	Q5	Q6	Total	Started	Duration	Status	Options
3636995			2.0	1.0	1.0	2.0	1.0	2.0	**9.0**	2006 Mar 09 09:32	50 min	Marked	view
3639045			2.0	0.0	1.0	2.0	1.0	2.0	**8.0**	2006 Mar 09 10:18	59 min	Marked	view
3611210			2.0	0.0	0.0	1.0	1.0	2.0	**6.0**	2006 Mar 09 09:25	58 min	Marked	view
3663379			1.0	1.0	1.0	1.0	0.0	2.0	**6.0**	2006 Mar 09 11:02	48 min	Marked	view
3769747			2.0	1.0	1.0	2.0	1.0	2.0	**9.0**	2006 Mar 09 10:15	38 min	Marked	view
3780663			1.0	1.0	1.0	1.0	0.0	2.0	**6.0**	2006 Mar 09 11:04	41 min	Marked	view
3931943			1.0	1.0	1.0	1.0	0.0	2.0	**6.0**	2006 Mar 09 10:59	44 min	Marked	view
3938977			2.0	1.0	1.0	2.0	1.0	2.0	**9.0**	2006 Mar 09 11:06	51 min	Marked	view
3977802			1.0	1.0	1.0	2.0	1.0	2.0	**8.0**	2006 Mar 09 09:41	58 min	Marked	view
3946256			1.0	1.0	0.0	1.0	0.0	1.0	**4.0**	2006 Mar 09 08:45	59 min	Marked	view
3935047			1.0	1.0	1.0	2.0	0.0	2.0	**7.0**	2006 Mar 09 09:07	60 min	Marked	view
3955056			1.0	1.0	0.0	1.0	1.0	2.0	**6.0**	2006 Mar 09 09:46	56 min	Marked	view

Figure A-5-4: Part of the display gained by following the link ‘Assignment One (View Marks)’ shown in Figure A-5-2. Names and some ID information have been removed to preserve student privacy.

Question 4

Part	Guess	Answer	Tolerance	Score	Comment
1	-120	112.9412	1.0%	0	Incorrect
2	3.4704	3.4706	1.0%	1.0	Correct

For the circuit shown below,



(i) Calculate the voltage at point A.

-120 V

(ii) Calculate the MAGNITUDE of the current through R1. (Enter only positive values)

3.4704 mA

Figure A-5-5: Part of the display gained by following one of the 'view' links shown in Figure A-5-4. For each student, the individualised assessment questions may be viewed, along with correct answers and submitted answers.

Home News Practice Assess Setup Logout

Oasis

Setup

Additional Statistics

Assignment One

Question	Class Mean	Std Dev	Maximum	Minimum
1	1.605	0.630	2.0	0.0
2	0.842	0.365	1.0	0.0
3	0.789	0.408	1.0	0.0
4	1.447	0.714	2.0	0.0
5	0.526	0.499	1.0	0.0
6	1.816	0.555	2.0	0.0

Figure A-5-6: Class statistics for the assignment displayed in Figure A-5-4.

Setup

Student Records

Search for:

Student Details for Van Den Hurk, Aidan

NetID:	avan116
Student ID:	3984885
Given Name:	Aidan
Family Name:	Van Den Hurk
Enrolments:	ELECTENG208 MECHENG270

Assessments:

Course	Title	Started
ELECTENG101	Assignment One Repeat	2005 Sep 19 08:44
ELECTENG101	Assignment One	2005 Aug 26 13:51
ELECTENG101	Assignment Two	2005 Oct 17 16:38
ELECTENG202	Assignment One	2006 Mar 09 10:55
ELECTENG208	Assignment Two	2006 Mar 21 08:36

Practice:

Date	Practices
2005/08/24	40
2005/08/25	27
2005/09/16	1
2005/10/16	33
2005/10/27	18

Figure A-5- 7: The results of a search for a particular student. Links may be followed to display assessment data like that shown in Figure A-5-5. Scrolling displays more practice data. Some personal information has been removed to preserve student privacy.

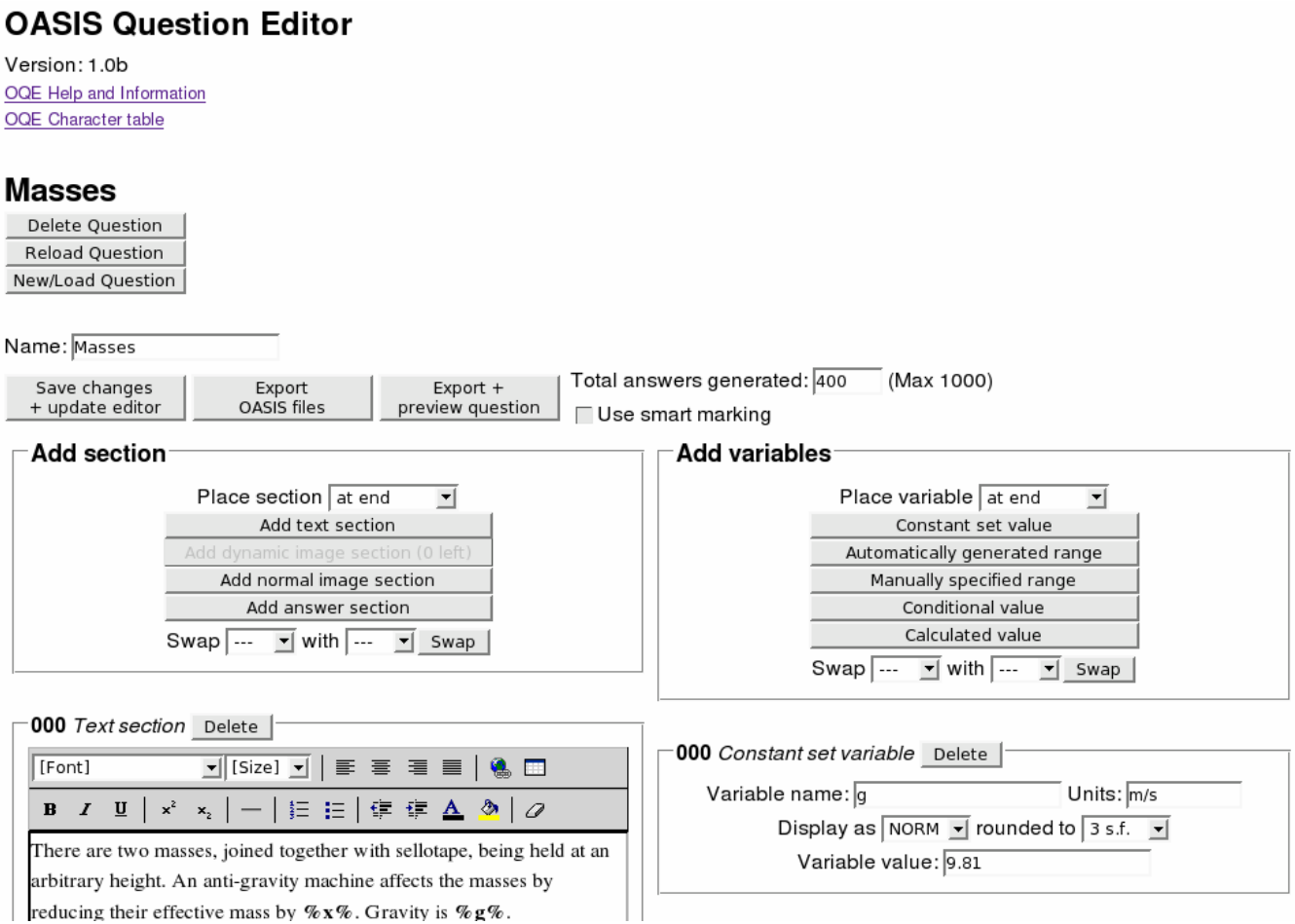


Figure A-5--8: The OASIS Question Editor.

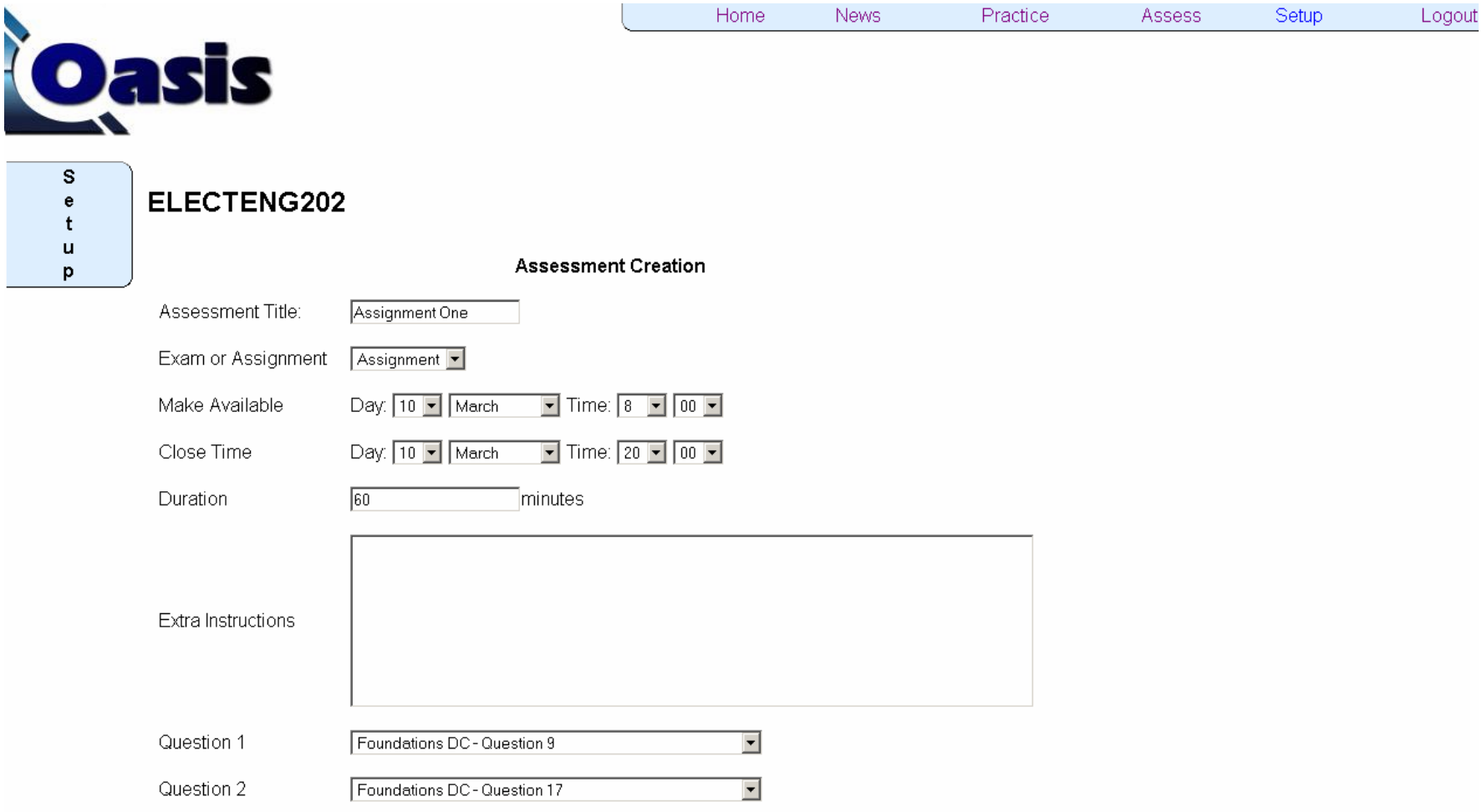


Figure A-5-9: The OASIS Assessment Creation page.

Appendix six: OASIS wish list 2002; staff access to database

This appendix contains the tables and notes I wrote in June 2002 to illustrate to my colleagues what features I saw as desirable for staff accessing the OASIS database.

Select paper

Electeng 101	Electrical engineering systems
Electeng 202	Circuits & systems

Select database

Electeng 202: Circuits and systems		
Go to student list (practice)	Go to question list (practice)	Go to assessment list

Student list (practice)

Electeng 202: Circuits and systems	
Aardvark, Aaron A.	Aaar001
Aardvark, Aaron B.	Aaar002

Question list (practice)

First select 100 series, 200 series, or 300 series, etc.
Then…

Electeng 202: Circuits and systems						
Question	Total student attempts	Total student time	Overall success rate	Number of 100% students	Average time for 100% attempt	Number of attempting students/ class size
101	435	28h 13m	84%	197	5m 27s	205 / 215
102						
103						
104						

Select a particular question to get more information…

Assessment list

Electeng 202: Circuits and systems	
Assignment 1 2 March	Test 1 23 March
Assignment 2 16 March	
Assignment 3 30 March	

Selected question (practice)

Paper 202
Question 105

Statement of question appears here

Total student attempts	357
Total student time	21h 05m
Overall success rate	83%
Number of students scoring 100% on one or more attempts	165
Average time for 100% attempt	4m 12s
Number of students attempting question	184
Class size	215

Selected student (practice)

Aardvark, Aaron A.
Aaar001
Paper 202

Logon date and duration	Questions attempted Success rate Time taken						
2 Mar 02 34m 28s	001 100% 3m 5s	002 100% 4m 17s	003 0% 5m 56s	003 50% 4m 23s	003 100% 4m 39s	004 60% 6m 28s	004 100% 5m 40s
4 Mar 02 47m 53s							

In above table, more information available about each question attempt by clicking on appropriate cell…

Selected question for student (practice)

Aardvark, Aaron A.
Aaar001
Paper 202
Question 105

Statement of question appears here

Statistics for class	
Total student attempts	357
Total student time	21h 05m
Overall success rate	83%
Number of students scoring 100% on one or more attempts	165
Average time for 100% attempt	4m 12s
Number of students attempting question	184
Class size	215

Statistics for student	
Total student attempts	5
Total student time	15m 3s
Overall success rate	80%
Number of times 100% scored	4
Average time for 100% attempt	2m 57s
Best time for 100% attempt	2m 01s
Ranking of best 100% time	43

Selected assessment

Electeng 202: Circuits and systems Assignment 3 30 March	
View by student	View by question

Assessment by student

Electeng 202: Circuits and systems Assignment 3 30 March						Export marks to Cecil	
		Total (15)	Q1 (2)	Q2 (3)	Q3 (3)	Q4 (3)	Q5 (4)
Mean mark							
Standard deviation							
Student	UPI						
Aardvark, Aaron A.	Aaar001	12	2	2.8	2	2	3

Click on question to view statement of that question
Click on student to get more detail about that student…

Assessment by student, detail

Electeng 202: Circuits and systems Aardvark, Aaron A. Aaar001 Assignment 3 30 March				
Question	Correct answer	Student answer(s)	Mark	Out of
Q1 (a)			1	1
Q1 (b)			1	1
Q2 (a)			1	1
Q2 (b)			1	1
Q2 (c)	9.2	18.4 9.2	0.8	1
Q3(a)			0	1

Click on question to view statement of that question

Assessment by question

Electeng 202: Circuits and systems Assignment 3 30 March							
	Q1		Q2			Q3	
	a	b	a	b	c	a	
First attempts correct, %	85%						
Second attempts correct, %	60%						
Maximum possible score	1						
Average score	0.922						

Appendix seven: SPSS analysis of OASIS practice data

Regression

Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	OASIS practice_total(a)	.	Enter

a All requested variables entered.
b Dependent Variable: EESexam_Total

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.192(a)	.037	.034	17.760

a Predictors: (Constant), OASIS practice_total

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3932.708	1	3932.708	12.468	.000(a)
	Residual	102509.147	325	315.413		
	Total	106441.855	326			

a Predictors: (Constant), OASIS practice_total
b Dependent Variable: EESexam_Total

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	57.716	1.833		31.491	.000	54.11	61.32
	OASIS practice_ total	5.575E-2	.016	.192	3.531	.000	.025	.087

a Dependent Variable: EESexam_Total

Regression

Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	UEBS Total, OASIS practice_total(a)	.	Enter

a All requested variables entered.
b Dependent Variable: EESexam_Total

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.610(a)	.372	.369	14.359

a Predictors: (Constant), UEBS Total, OASIS practice_total

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	39638.978	2	19819.489	96.126	.000(a)
	Residual	66802.877	324	206.182		
	Total	106441.855	326			

a Predictors: (Constant), UEBS Total, OASIS practice_total
b Dependent Variable: EESexam_Total

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-26.721	6.585		-4.058	.000	-39.67	-13.766
	OASIS practice_ total	4.121E-2	.013	.142	3.216	.001	.016	.066
	UEBS Total	.239	.018	.581	13.16	.000	.203	.275

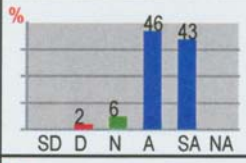
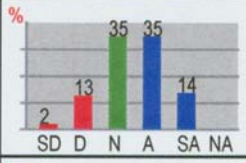
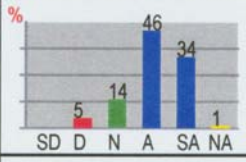
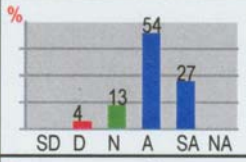
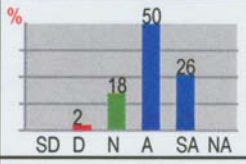
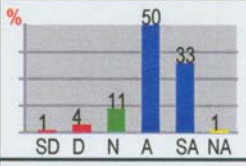
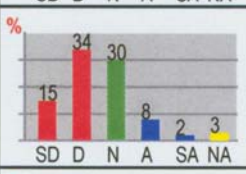
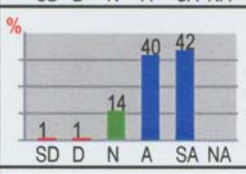
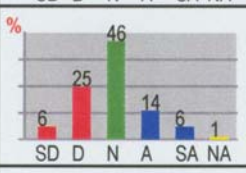
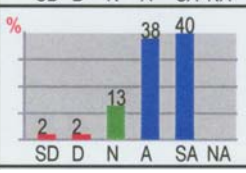
a Dependent Variable: EESexam_Total

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
EESexam_Total	327	13	99	63.18	18.070
OASIS practice_total	327	1	296	98.01	62.303
UEBS Total	327	218	464	359.12	43.936
Valid N (listwise)	327				

Appendix eight: Survey analysis, 'Circuits & Systems', 2004

M138 OASIS is easy to use		A+SA = 89% D+SD = 2% count=3
M139 I often use OASIS		A+SA = 49% D+SD = 15% count=3
M072 OASIS helped improve my skill level		A+SA = 80% D+SD = 5% count=3
M140 I am more confident about my learning after using OASIS		A+SA = 81% D+SD = 4% count=3
M126 The problems provided in OASIS helped me understand the course material better		A+SA = 76% D+SD = 2% count=3
M141 OASIS helps me to prepare for the assessments		A+SA = 83% D+SD = 5% count=3
M142 I come to campus less often because of OASIS		A+SA = 10% D+SD = 49% count=3
M083 I like the instant performance feedback using OASIS		A+SA = 82% D+SD = 2% count=3
M143 I prefer to do problems from the text rather than OASIS problems		A+SA = 20% D+SD = 31% count=3
M144 It would be a good idea to have OASIS in other courses		A+SA = 78% D+SD = 4% count=3

Appendix nine: Publications list associated with OASIS

Journal articles

C. Smaill, “The implementation and evaluation of OASIS, a web-based learning and assessment tool for large classes,” IEEE Transactions on Education, vol. 48, no. 3, pp. 658-663, November 2005.

S. Hussmann and C. Smaill, "The use of web-based learning and communication tools in electrical engineering," Australasian Journal of Engineering Education, vol. online publication 2003-01, 2003. Available at http://www.aaee.com.au/journal/2003/hussmann03.pdf

Conference Presentations

C. Smaill and S. Hussmann, “The implementation and evaluation of an individualised, web-based, formative and summative assessment tool for large classes,” Proc. of 4th International Conference on Information Technology Based Higher Education and training, Marrakech, Morocco, July 2003, pp. 87-92

A. Bigdeli, J. Boys, C. Smaill, and C. Coghill, “Can a computer-assisted marking system measure student’s understanding?”, Proc. of IEEE conference on Artificial Intelligence Systems, Divnomorskoe, Russia, September. 2003, pp. 116-124

C. Smaill, “The implementation and evaluation of a web-based learning and assessment tool in Electrical Engineering,” Proc. of Association for Engineering Education in Southeast and East Asia and the Pacific Conference, Auckland, New Zealand, December 2004, pp. 71-76

C. Smaill, “Student voices in computer-assisted learning,” presented at Society for Teaching and Learning in Higher Education Annual Conference, Charlottetown, Prince Edward Island, Canada, June 2005.

C. Smaill, “The implementation and evaluation of OASIS, a Web-based learning and assessment tool in electrical engineering,” Proc. of ASEE/IEEE Frontiers in Education Conference, San Diego, CA, USA, October 2006.

C. Smaill, "The student response to OASIS, a Web-based learning and assessment tool in electrical engineering", presented at 17th AAEE Annual Conference, Auckland, New Zealand, December 2006.

Appendix ten: Abbreviations

AAEE	Australasian Association for Engineering Education
ASEE	American Society for Engineering Education
CAA	Computer-assisted assessment
CAL	Computer-assisted learning
CAPA	Computer-Assisted Personalised Approach, *also* Computer-Assisted Personalised Assignments
CBA	Computer-based assessment
CML	Computer-managed learning
CPD	Centre for Professional Development
ECE	The Department of Electrical and Computer Engineering at UoA
EES	Electrical Engineering Systems
ERIC	Education Resources Information Center
IEEE	Institute of Electrical and Electronics Engineers
LON-CAPA	Learning Online Network with Computer-Assisted Personalised Approach
MSU	Michigan State University
OASIS	Online Assessment System with Integrated Study
PDA	Personal digital assistant
PPA	Paper-and-pencil assessment
SPSS	Statistical Package for the Social Sciences
UEBS	University Entrance and Bursaries and Scholarship Examinations
UoA	The University of Auckland
UWA	The University of Western Australia