University of Leeds

SCHOOL OF COMPUTER STUDIES

RESEARCH REPORT SERIES

Report 95.28

Open Book Assessment in Computing Degree Programmes[1]

by

Tony Jenkins

Division of Operational Research and Information Systems

September 1995

Abstract

The use of "Open Book" assessment is widely seen as an approach that more effectively tests a student's understanding and knowledge of a subject area. There have been a number of attempts to introduce such assessment in the School of Computer Studies at the University of Leeds, and open book examinations are now used in several modules.

This paper describes the experience of introducing open book assessment into a level 2 computing module. We consider the process of setting and marking the examination, and the problems and issues raised by the students. We also compare the results from the examination with the results of the same students on other modules examined "traditionally".

We demonstrate that open book assessment produces results that are at odds with the results produced from conventional assessment, and consider some of the reasons for this. It is also shown that students' results across open book examinations show a much weaker correlation than might be expected.

On the whole, open book assessment is shown to be a valid and useful way of assessing student learning. We conclude by considering whether it should be used to *replace* or *complement* conventional examinations, and what other assessment techniques may be used.

1 Introduction

In the School of Computer Studies at the University of Leeds there are two conventional ways in which modules in a degree programme are assessed. Modules that aim to teach essentially practical skills (such as programming or software engineering) are assessed entirely using coursework exercises. These modules are in the minority; the majority of modules are assessed with a "traditional" examination, perhaps with an element of coursework. This paper describes the experience of replacing this "traditional" examination with an "open book" examination in a second level module.

Section 2 begins by briefly describing the generally accepted theory of learning and the programmes and modules in the School of Computer Studies. In section 3 we consider the motivation for making a change to open book examination assessment, and the experience of the examination is described, and the effectiveness of the change from the point of view of both staff and students is discussed. An analysis of the results from the open book examination compared to the results of the same students on other modules is presented in section 4, and the issues raised by this are considered in more detail in section 5. The conclusion in section 6 summarises the experiences and describes how this work will be taken forward.

2 Background

This section describes the issues that were the basis for the introduction of open book assessment in a particular module run by the School of Computer Studies at the University

of Leeds. The necessary background concerning the degree programmes in the School is also briefly described.

2.1 What do Examinations Assess?

The purpose of any examination (or indeed any form of assessment) is presumably to assess the extent of a student's *learning*. When a conventional examination (usually held in intimidating and unnatural conditions) is used this may not be the case. In [1] Phil Race and Sally Brown present a consideration of what examinations *really* measure. They propose this list as an agenda for discussion:

- How much you know?

- How much you don't know?

- How fast you can write?

- How good your memory is?

- How much work you did the night before?

- How well you kept your cool?

- How practised you are?

- How well you can read the questions?

- How good you are at answering exam questions?

- How *practised* you are at answering exam questions?

- How you perform under pressure?

- How good you are at time management?

- How well you can keep addressing the question?

- How *often* you've practised on similar questions?

- How well you read your own answers?

The most striking thing about this list is the absence of the question "How much you have learned?". Apart from the first two questions all the issues raised relate to *exam technique*, and practice in taking examinations. It is clear that much of what an examination assesses, therefore, is a student's ability to take examinations; the student's knowledge (or learning) of the subject area is secondary. Anyone who has set or taken a traditional examination will recognise the truth in this.

If we accept, therefore, that an examination is an inefficient (and possibly ineffective) way of assessing a student we must try to devise a replacement. The most obvious solution is to replace the examination with some form of continuous assessment. This is an appealing idea,

but it brings many problems of its own. Foremost among these are plagiarism, and the extra demands placed on staff in terms of setting and monitoring assessments. Especially with the current increases in class sizes in Higher Education it is unlikely that this will become a feasible approach in the near future.

The requirement, then, is to devise a method of assessment that places less emphasis on a student's examination technique, but still retains the efficiency of a formal examination. Thus the focus is firmly on the "What do you know?" issue. Before devising such an examination, however, we need to define what is meant by *learning*.

2.2   Theory of Learning

It is generally accepted that learning can be described as *surface* or *deep* (there is a good account in [2]). Surface learning is generally learning by rote of what is perceived by the student as important. The students focus on isolated facts and usually fail to see the overall relationships between the material. On the other hand deep learning represents an appreciation of these overall relationships. We can reasonably summarise surface learning as *memorising* and deep learning as *understanding*.

David Baume [3] refines the deep and surface classification to distinguish between the learning of *information* and the learning of *meaning*. Students who adopt a surface learning strategy generally learn information or *facts* whereas deep learners attempt to learn the underlying meaning of the material they are being taught.

It seems obvious that examinations should test deep learning (and the understanding of meaning) rather than surface learning (the memorising of facts). With many traditional examinations it would be possible for a student to pass (and pass well) solely as the result of surface learning. Indeed, it will be shown later how, in some cases, a deep learner might actually be disadvantaged in certain types of examination.

2.3   Open Book Examinations

In a conventional examination the student is generally required to answer the questions with no outside help. Any help or reference material is usually limited to a formula sheet or set of tables. In an open book examination, the students are allowed to use any reference materials which they have brought with them. Apart from this, normal examination regulations apply; there is a time limit, and no communication with other students is allowed.

The motivation for the use of such examinations is simple. If the students have reference materials with them, there is no point in surface learning; all the material that might have been learned by rote is available to all the students during the examination. Thus open book examinations may test *deep* learning.

Traditional examinations tend to pose questions that encourage surface learning [4], and so it is reasonable to expect students to adopt such strategies. To take an extreme example, a traditional examination question might be:

List the fifteen factors affecting student motivation identified by Henderson.

Clearly, to gain full marks in this question a student must have **memorised** the fifteen points. The student does not need to understand the issues; all that is being assessed is memory. Moreover, a student who has a deep understanding of the points made by Henderson may find themselves at a disadvantage as they cannot remember the precise fifteen points, or perhaps the division between them.

In an open book examination, on the other hand, questions can be set that test **understanding**. For example:

Discuss the validity of the factors identified by Henderson in relation to your studies at this University.

Here it is plain that surface learners will be at a disadvantage. They will probably be able to gain some credit but will have to work "on the fly". Deep learners should be able to answer the question more fully, and should thus achieve better results. Notice that the open book question does not mention the subject of Henderson's factors; a well-prepared student will know the subject area and will be able to find the relevant notes quickly.

Another issue, that will be considered later, is that the marking scheme for the open book question must be different. In the first question, it would be simple to allocate one mark to each (correct) point given in the answer. This cannot be done for the second question; a marking scheme that assesses the student's *learning* must be devised.

These suggested advantages for open book examinations focus on the learning experience of the students. While this is the most important issue, it will be shown later how the use of open book examinations can be more interesting, challenging and satisfying for the assessor. It can at times even be entertaining!

2.4   Degree Programmes

The School of Computer Studies at the University of Leeds offers two single subject degree programmes, Computer Science and Information Systems. These programmes are taught as a set of *modules* taken over three years. In addition, students from other departments may take modules in the School, and a number of exchange students arrive every year. The modules may be a compulsory component of a degree programme, or they may be taken as options or electives. An elective is a special type of option that is intended to encourage students to take modules outside their main subject areas; such modules do not usually count towards a final degree classification.

There are approximately 350 single subject students in the School and some 600 students from outside. The average class size is about 120, although there is a very wide variation.

The year is split into two semesters, and there is an examination period at the end of each. The majority of modules have an examination (usually worth 80% of the assessment) at the end of the semester in which they are taught.

3 The Examination

This section describes the practical motivation for changing the assessment for a particular module, and then describes the process of setting and sitting the paper. It concludes by considering the feedback from the staff and students concerned.

3.1 The Module

The module concerned here is entitled *Commercial Systems Development*. The School code for the module is *SE4*, which will be used here. It is available as an option or elective only, and is normally taken by level 2 (second year) students. The class size is usually between 100 and 120.

The aim of the module is to give the students an appreciation of the application of systems development and software engineering in the "real world". This is supported with practical work using the Ingres database and fourth generation language. Other topics include CASE, prototyping, open systems, and the problems inherent in developing a commercial information system.

3.2 Motivation for Change

SE4 was the author's first lecturing assignment starting in the academic year 1993/94. In this year the module was taught largely as it had been before; it was assessed mainly with a traditional examination.

When marking this examination, it became obvious that many students were gaining reasonable marks simply for having memorised the course notes. Moreover there were some students who clearly understood the material well, but were obtaining relatively low marks because they failed to mention key points or use key words. With the benefit of hindsight it is apparent that the examination was testing *surface* rather than *deep* learning.

At the same time, colleagues in the School were experimenting with open book exams. Dr Roger Boyle and Dr Mourad Kara had both introduced these exams into their level 2 modules and seemed happy with the results. It was therefore decided to introduce an open book examination into SE4.

The final motivation for change was rather more selfish. Marking an examination where the majority of the students are presenting the same answers in the same words was an extremely tedious task. It seemed likely that an open book exam might also provide more interesting marking.

3.3   Setting the Examination

Traditionally the examination for SE4 had consisted of four (the students being required to answer three) straightforward questions such as:

Describe, with the aid of a diagram, the traditional Systems development life cycle.

[7 Marks]

These questions required the students to have memorised a particular part of the course notes. The later parts of the questions often required more understanding (deep learning), but even here it was possible to achieve a reasonable mark simply by reproducing the course notes. These questions were thus "easy" to answer and straightforward, but tedious, to mark.

It was anticipated that the setting of the open book examination would be harder than the setting of the traditional paper. This proved not to be the case; in many respects it was easier. The approach was to set questions that gave the students the opportunity to show their *understanding* of the material covered in the module or to explore their own views on an issue. For example, the following question was used in 1994/95:

In Commercial Systems Development the human-related problems (at both a personal and organisational level) are often more important, and harder to overcome, than the purely technical issues.

Discuss the validity of this statement in relation to **FOUR** *of the following:*

- *Problems with the traditional Systems Development Life Cycle;*

- *End-User Computing;*

- *Systems development using a Prototyping Methodology;*

- *Information Centres;*

- *A long term systems development project where the customer site is a large distance from the base of the systems developers.*

[5 Marks Each]

This question was intended to allow the students to develop their own ideas on these issues; the issues themselves had not been covered explicitly in the lectures.

A set of questions was devised quite quickly. An effective approach was to include a small case study in a question and to ask the student to analyse the issues involved. A copy of the paper finally used is included as an appendix to this paper.

Unfortunately, even though setting the paper was reasonably easy, there were problems in producing a marking scheme. A marking scheme is required that shows clearly where each mark for a question will be available. For a traditional examination question such as that

6

quoted above this is straightforward. However, in an open book examination it is much harder to forecast how the students will answer the question, and so it is very difficult to allocate marks to specific points. Indeed, in many cases there was no "correct" answer to a question. Eventually, it was decided to allocate marks to be awarded based on an assessment of the student's understanding. It remained to be seen whether this would be satisfactory.

The final change from the traditional paper was to change the examination rubric to indicate the open book nature of the exam. The wording finally chosen was:

Candidates are permitted to bring with them any notes (hand-written or computer-generated) that they may have. Photocopies of other material are also permitted.

Books, other than dictionaries, are not permitted.

Calculators are permitted.

You are reminded that, in normal circumstances, there are no marks available for the direct reproduction of the lecture notes.

Books were banned so that students who did not have (or were unable to afford) certain books were not disadvantaged. It was also hoped to stop the practice of hiding or stealing library copies of these books. The rubric also contained a specific warning against reproducing the course notes; the effectiveness of this warning will be considered later.

3.4   Sitting the Examination

The vast majority of the students enrolled for SE4 had never taken an open book exam before and so it was necessary to prepare them. This preparation had the useful side effect of allowing the "prototyping" of an examination question in order to assess what the students might reasonably be expected to do.

An example question was prepared and was discussed in a lecture. Some initial feedback was gained during the lecture and the students were invited to submit individual comments later. This session had an interesting result; no students offered any feedback, but several quickly changed SE4 to an elective (and thus a module that did not count towards their degree), and a few even dropped the module.

For the examination itself, there was a wide variety of approaches apparent. A few students had just a copy of the lecture notes; one had nothing at all. On the other hand several brought huge amounts of material with them; some clearly had several photocopies of complete books. The average seemed to be to have the course notes with hand-written annotations.

3.5   Marking the Examination

As expected, the marking proved to be more interesting than the marking of the previous paper. Some of the answers were extremely well written and were almost entertaining. However, it quickly became clear that the marking would take much longer than the marking of the previous exam since all the answers had to be read in detail, and sometimes even considered at some length.

An interesting feature was that some students proposed ideas or points of view that were totally opposed to those presented in the lectures. There seemed no reason why they should not gain credit for these ideas and, where they were suitably justified, they were given equal credit to that available for the "establishment view".

It was also apparent that some students had totally failed to appreciate what was expected of them. A large number still reproduced their notes directly and so achieved very low marks. This seemed to be the result of "keyword spotting"; the student saw a word in the question and immediately copied out the relevant section of their notes. This attitude is probably caused by the unnatural and intimidating way in which examinations are still sat.

3.6 Student Feedback

The School of Computer Studies provides a bulletin board [5] on which students are encouraged to post their comments or questions on a module. It was anticipated that the SE4 examination would provoke some comments here, and this proved to be the case. The articles in this section are reproduced verbatim.

The first article was from a student who was highly critical of the examination:

```
Newsgroups: local.modules.se4.talk
From: csynsc@scs.leeds.ac.uk (N S Curzon)
Subject: the exam
Date: Thu, 2 Feb 1995 12:19:38 GMT

The exam (or masscre in my case),

What can I say other than what was the point in making it an open book exam?
None what so ever! Tony's notes were next to useless, the previous years exam
paper was a waist of space and of paper, any any photocopies and extra notes
that I made didn't help much either. Now, I except some of the blame, that I
photocopied the wrong parts from a book and so on. But I feel that I have
waisted my time, that I would have been better off with an ordinary exam that
required me to learn the notes. At least then it would not seem so pointless.

Overall, this model has been a waist of time (for me anyway), and I wish I had
seen this sooner. I would then have had the oportunity to do something that was
actually useful.


--
Neil Curzon
```

This was clearly somewhat depressing. The first reply, however, is very interesting and was at the time very reassuring[2]:

```
Newsgroups: local.modules.se4.talk
```

---

[2]The quoted article has been removed here.

From: csycmb@scs.leeds.ac.uk (C M Barnes)
Subject: Re: the exam
Date: Thu, 2 Feb 1995 12:44:29 GMT

But wasn't the idea of the module to make you think about how things
actually happen in the real world etc. There is more to life than programming.
I'm glad it was an open book exam. I certainly couldn't have remebered enough
to do the exam. Tonys notes weren't much good if thats all
you had as they just seem lists of points. However, if you had added to the
points in the lectures they actually became quite usefull. I suspect that Tony
had done this deliberatly.

--
Catharine

These articles were followed by a long debate about the merits of the examination. The
general opinion seemed to be in favour of the open book exam. Two interesting extracts
were:

Newsgroups: local.modules.se4.talk
From: csypgw@csgi23 (P G Windle)
Subject: Re: the exam
Summary: Pete leaps to SE4's defence, in the hope of bonus marks on the exam :)
Date: Thu, 2 Feb 1995 13:02:35 GMT

I thought that the exam was a good measure of how well you actually understood
the issues involved, rather than your ability to find relevant reference texts
and pay 7p/A4 sheet.

: Tony's notes were next to useless,

Not so. Instead of memorizing the low-level details, my revision time (all 4
hours of it) was spent trying to get my noggin around the higher level concepts
involved in large scale systems development. During the exam, should I need to
reference some list of features or reasons, I just flicked through said notes.

In this case, SE4 is a module which either consists of rote learning of
lists which are variable anyway - said list of 4GL attributes is only an
opinion - or use of facts in the context of CSD to discuss the issues
constructively. Use meta-knowledge of the subject if you will.

The thing is, in the real world, no-one takes away your reference texts. If
someone were to say: "Write a report on subject X. You have a month to read up
on the subject. To produce your report, you will be locked in a room without
notes. You will have pen and paper; no other resources will be permitted."
you would think that they were insane. In this module, we were expected to

```
acquire knowledge, not to memorize facts. The exam tested how we could apply
the knowledge gained to problems.

I feel that the SE4 exam was a true reflection of what an open-book exam
should be like. In fact, although I may not have attended many lectures, I
found the entire module to be the most enjoyable yet, bar none.


Newsgroups: local.modules.se4.talk
From: isydps@scs.leeds.ac.uk (D P Sugden)
Subject: Exam
Date: Thu, 2 Feb 1995 14:43:23 GMT

Well I hate to be agianst what everyone seems to have written BUT

I actually enjoyed the exam!!! Yes hard to believe, and it is the first
one that I have really got into for many years.
Now I am not saying that I passed or anything like that, becsause some
of my answers were a tad weird admittedly but at least it was interesting,
and to me it seemed more "real".
Open Book was good.
```

The general view of the students seemed to be that the open book examination had been a
success. The response of some students (such as the first reproduced here) was disappointing,
but this was far from the majority view. It was especially pleasing that some of the students
have clearly understood the reasons for the introduction of the open book examination, and
that they approve of these reasons. These issues were never formally discussed in lectures.

3.7   Student Performance

As will be shown in more detail later, the marks for SE4 showed much the same spread of
marks as other modules taken by the same students. There are two interesting cases that
are worthy of note at this point, however.

One student passed SE4 with a grade of 53. This is remarkable because he failed every other
module with a grade of 25 or less. Clearly he had not done the necessary learning to pass
the other modules, but the open book approach gave him the opportunity to show his actual
level of ability. In another case, a student who was very worried that he had not produced
"correct" answers (the student in the last extract above) scored the highest mark of all on
the exam.

3.8   Conclusions

On the whole the use of an open book examination in SE4 was a success. It is especially
pleasing that the students themselves appear to be largely in favour of the idea. From the

point of view of the assessor, there was the satisfaction that the *understanding* of the students had been assessed rather than the memory.

When the SE4 results were compared with the results of the same students on different modules, some striking differences were apparent; a student who was routinely scoring first class marks had not done so in SE4, and there were cases of students failing other modules and doing very well in SE4. It was therefore decided to carry out a more formal comparison of the results; the results of this analysis are presented in the next section.

4 Analysis

In this section, the results of the SE4 are examined in the wider context. The results achieved by the students are compared to the results achieved by the same students on other modules.

This analysis is necessary in order to establish the validity of using an open book approach when the main assessment in other modules is carried out in a different way. The results have an influence on whether open book examinations can or should be more widely used.

4.1 Other Modules

Before comparing the results, a brief explanation of the other modules is necessary. They are as follows:

HC1 - *Introduction to Human Computer Interaction.* This module deals with issues such as interface and dialogue design. It is assessed by coursework (60%) and an open book examination (40%).

IS3 - *Introduction to Database Principles and Practice.* A general introduction to database issues such as distribution and query optimisation and processing, IS3 builds on a more basic introduction given at level 1. IS3 is assessed mainly by a traditional exam (worth 80%) and coursework.

IS4 - *Information Systems Design.* IS4 is a general introduction to the issues involved in the design of an information system. The issues covered include file structures, design audit, and personnel issues. It is a natural precursor to SE4. Assessment is by traditional exam (80%) with the remainder from coursework.

IS3 and IS4 are used because they cover the material that is closest to that covered in SE4. The three modules are often taken together, although this is not always the case. HC1 is included as it is the only comparable module that is assessed with an open book exam.

In the following sections, the analysis ignores those students that take either module as an *elective.* Since electives do not count towards a final assessment many students aim merely to pass, and concentrate their efforts on other modules. These students are ignored so that they do not distort the figures.

11

4.2   Traditional Examinations

If it is accepted that traditional examinations test surface learning, it is reasonable to suppose that a surface learner can perform equally well in any examination, irrespective of the subject material. This hypothesis was tested by plotting a scatter chart of the results of students taking IS3 and IS4 (figure 1). This figure clearly shows the expected trend.
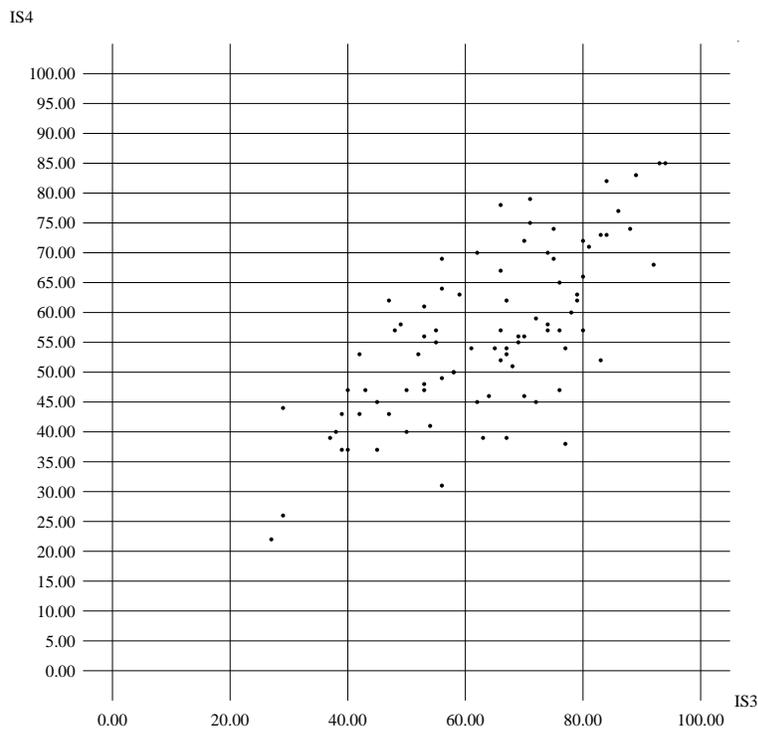

Figure 1: IS3 and IS4 Results

In order to make quantitative comparisons, a correlation value[3] was calculated for the IS3 and IS4 marks. The correlation is 0.71 based on a sample of 86 students. This value is high enough to suggest that there is a significant correlation between student performance on the two modules.

4.3   Traditional and Open Book Examinations

The next stage is to compare SE4 with a traditional examination. The chart is shown in figure 2; the correlation is 0.46. This suggests a correlation between the results but, importantly, it is a much weaker correlation than that found between IS3 and IS4.

Again this is what might be expected. Since the SE4 and IS4 examinations are testing different types of learning, the correlation should be lower. However, there is still some correlation, probably because the material in the two modules is closely linked. Indeed, some students reported using material from SE4 to answer parts of the IS4 examination.

---

[3]Pearson's correlation was used. This gives a value between -1 and 1 where 1 represents total positive correlation, -1 total negative correlation and 0 no correlation.
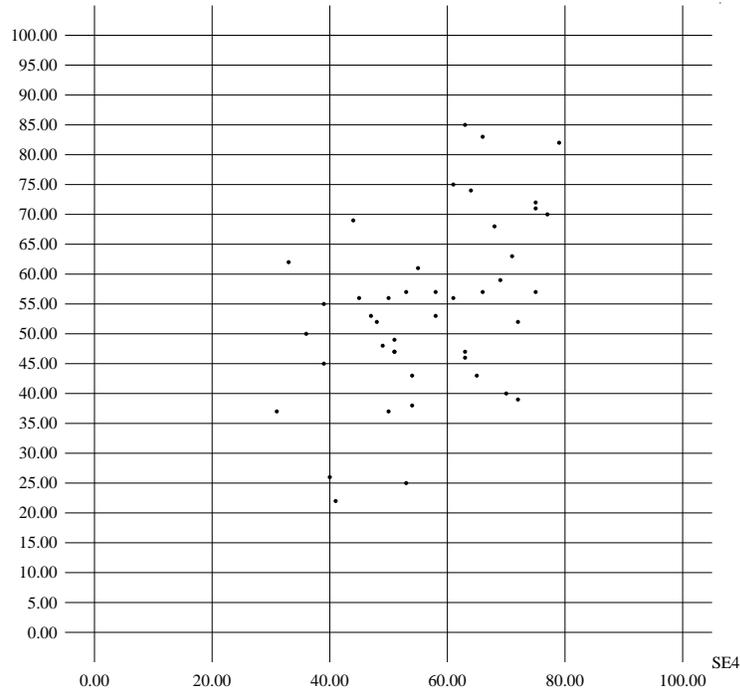
Figure 2: SE4 and IS4 Results

4.4 Open Book Examinations

To examine the relationship between the marks obtained in two open book examinations, the results from SE4 and HC1 were compared (figure 3). This shows some correlation (the value is 0.40), but it is much lower than that found between two traditional examinations. An interesting feature of the chart is the apparent "clumping" of the results; they are much less spread out than in, for example, figure 1.

It might be expected that two open book examinations would have a higher correlation, but this proves not to be the case. The reason for this is that the open book examinations are successfully testing a student's deep learning in a subject area. With these two modules the subject areas are fairly separate; there is no reason to suppose that a student who has a good understanding of human-computer interaction issues has an equally good understanding of the issues covered in SE4. The correlation that exists shows that there are some students who perform equally in both, but that these are not the majority.

This result was verified by removing the coursework components from the figures (HC1 has 60% coursework and SE4 20%) in case the coursework marks were affecting the figures. The revised chart is shown in figure 4. The correlation in this chart is slightly higher (0.46), but still significantly lower than that found between the two traditional examinations.
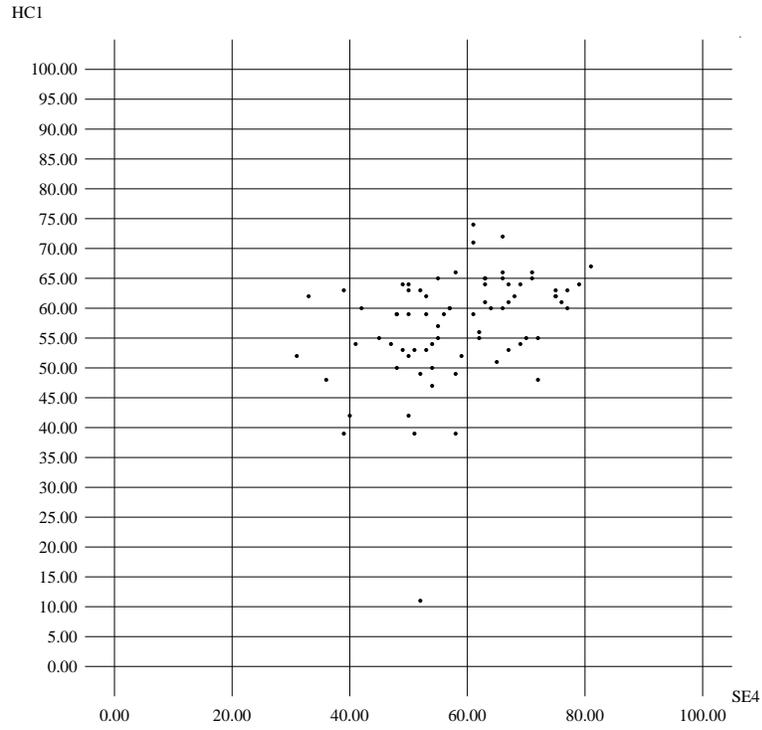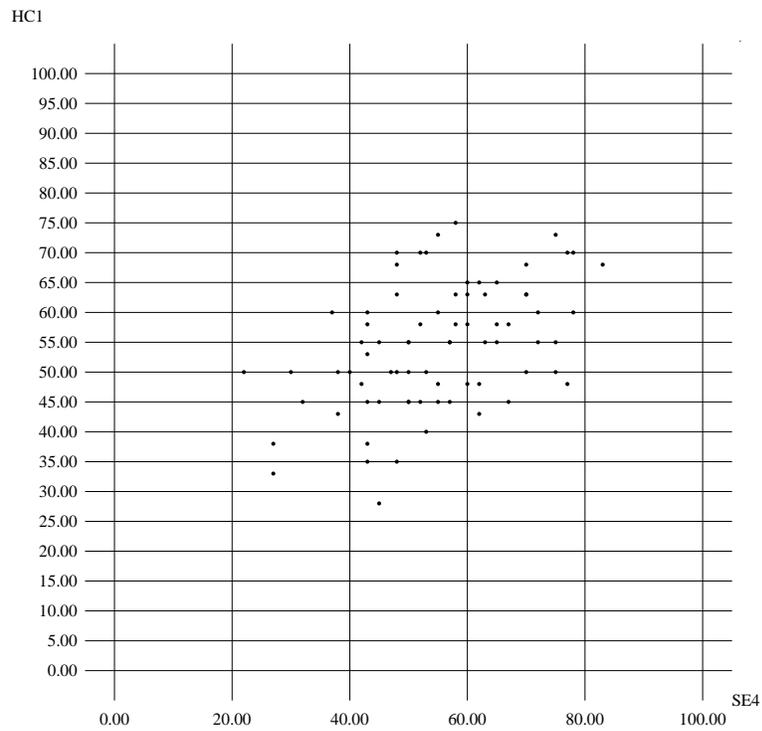
Figure 3: SE4 and HC1 Results

Figure 4: SE4 and HC1 Results (Excluding Coursework)

14

4.5 Further Analysis

In the light of these results, it was decided to compare SE4 and IS3 with a module covering material that was as far removed from SE4 and IS3 as possible. The module chosen was that which gave the largest sample size, *TC2 - Theory of Computation*. This is a highly theoretical module dealing with topics such as automata theory, formal language theory, and the theory of abstract data types.
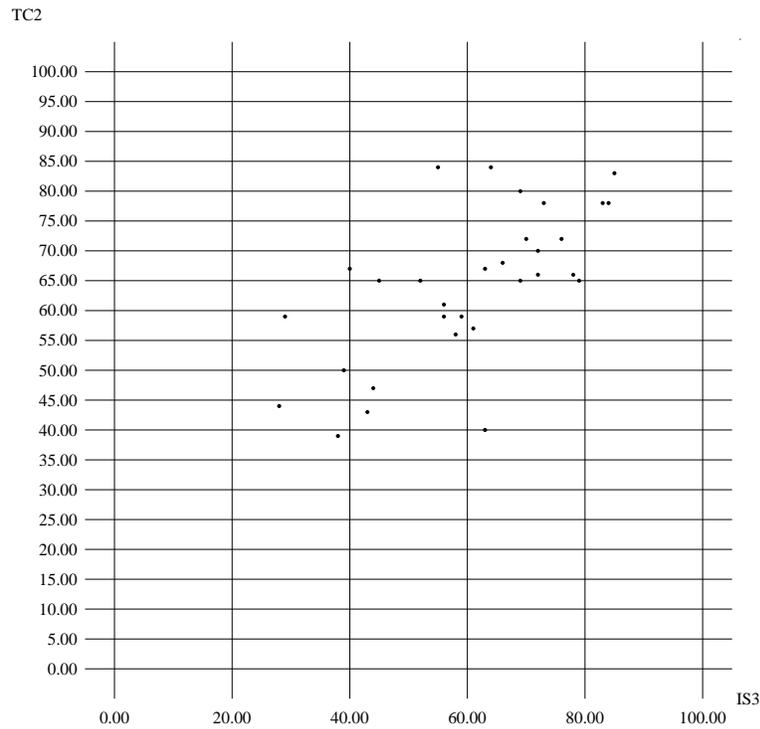

Figure 5: IS3 and TC2 Results

The results of the comparison are shown in figure 5 and 6; the correlations are 0.65 for IS3 and -0.04 for SE4. Thus there is a reasonably strong correlation between the results of IS3 and TC2, where there is no strong overlap in material; students can pass using surface learning alone. Furthermore, there is no correlation at all between SE4 and TC2; this seems reasonable as SE4 is a discursive module dealing with systems development and software engineering whereas TC2 is highly theoretical. It seems quite likely that students who do well in a theoretical module such as TC2 would not do as well in the more discursive SE4.

A comparison between IS4 and TC2 was rejected due to the small number of students involved (5). The correlation here is actually higher, 0.74. As a further control TC2 was compared with another more theoretical module, *AI4 - Computer Vision*. This gave a correlation of 0.89, which is as expected for two traditionally examined modules.
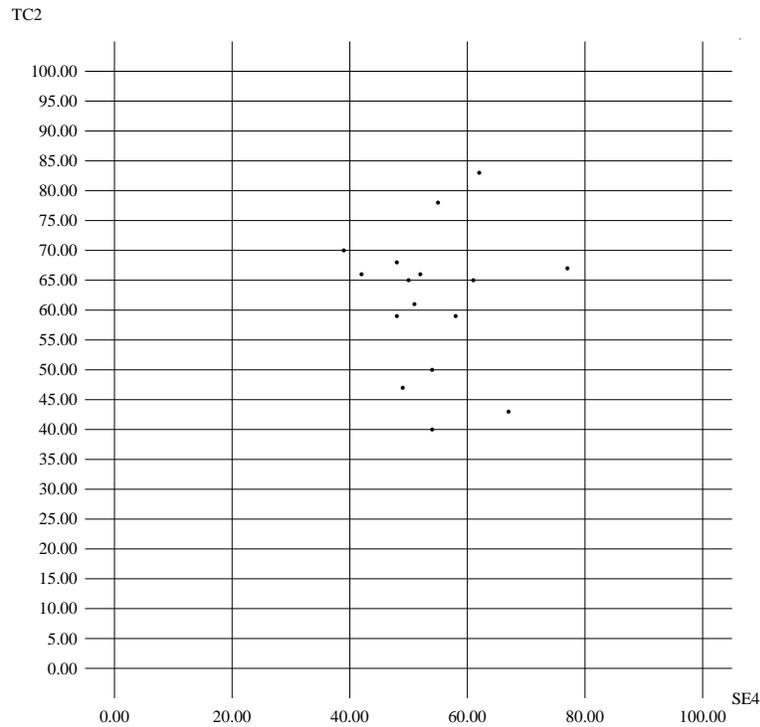
Figure 6: SE4 and TC2 Results

4.6 Summary

The correlations calculated are as follows:

Module	Module	Sample	Correlation	Comments
IS3	IS4	86	0.7125	
SE4	IS4	44	0.4598	
SE4	HC1	74	0.4029	
SE4	HC1	74	0.4629	Excluding Coursework
IS3	TC2	31	0.6498	
IS4	TC2	5	0.7354	Disregarded - Sample Size
SE4	TC2	16	-0.0396	
TC2	AI4	12	0.8877	

SE4 and HC1 have open book examinations, the other modules have traditional examinations.

4.7 Conclusions

This analysis clearly shows that there is a strong correlation between the performance of students in traditional examinations, and that this correlation is affected only slightly by the subject material concerned. This shows that surface learners are able to memorise

material on any subject; they have no need to understand it, so the subject material itself it unimportant.

On the other hand, the correlation reduces as soon as an open book examination is introduced. This is what would be expected; this examination is testing deep learning, the extent of which is much more closely related to personal interest and aptitude.

One crucial aspect of assessment is that it should be consistent. Especially in a modular system, this consistency must apply across all modules. The analysis presented here indicates that the introduction of open book testing (while traditional testing is still present) has an effect on the consistency of the assessment process. This effect appears to be potentially damaging. This issue is considered in section 5.

5 Changing to Open Book Examinations

Sections 3 and 4 have highlighted some of the problems that may be encountered when changing to open book assessment. This section considers these and other issues that arise when such a change is considered.

5.1 Student Performance

To assess the impact of a change to open book assessment an obvious analysis is to compare the average student performance before and after the change. The figures for SE4 are[4]:

Session	Examination	Sample	Average Mark
1993/94	Traditional	71	59.09
1994/95	Open Book	80	57.94

Thus the introduction of open book assessment has lead to a slight decrease in the average mark. However, this decrease is insignificant and is no more than might be expected between any two classes of students. To demonstrate this, the figures for IS4 over the same period are:

Session	Examination	Sample	Average Mark
1993/94	Traditional	66	54.26
1994/95	Traditional	91	55.46

Another aspect of student performance is the range of marks achieved by the class as a whole. The scatter charts show that it appears to be rather less common for students to score a low mark in an open book examination than in a traditional examination. More strikingly it is much more common for a student to score a very high mark on a traditional examination that in an open book examination (compare figures 5 (traditional) and 3 (open book)).

---

[4]Students taking the module as an elective are again excluded.

This can be explained by again considering what is begin tested. In a conventional examination a student, in general, either knows the answer to a question or they do not. A student who has memorised the entire syllabus will be able to achieve a very high mark; a student who has done no work will achieve a very low mark. In an open book examination, however, this does not hold. Since there is no "correct" answer, a student cannot simply reproduce it and thus score full marks. Similarly, a student who has not prepared in the conventional way may be able to achieve a reasonable mark simply by expressing their own views in the examination. The student feedback and results reported in section 3 demonstrate that this is so.

Whether this observation should affect a decision on whether to introduce open book examinations is open to debate. The advantages that have been described in terms of assessing students' *learning* rather than their memories should outweigh the consideration of consistency. In any case, there are so many more factors at work, that it cannot be said for certain that it is the open book assessment itself that affects the marks. For example, the lecturers that introduce this form of assessment tend to the more innovative and furthermore tend to use other innovative methods in presenting their material; it may be that these factors affect the results as much as the method of assessment.

5.2   Ensuring Fairness

It is essential that the assessment of a module is fair for all the students taking that module. Since open book examinations rely on the students developing their own ideas, normally in an essay form, it seems likely that students who have trouble with language will be at a disadvantage. This might include:

- Overseas students whose first language is not English;

- Students with a condition such as dyslexia;

- UK students with a low level of proficiency in English.

To address the problems of the first of these groups, the results for SE4 and HC1 were again analysed, this time splitting the UK and overseas students. The results are:

Module	Class Size	Average	UK	UK Average	OS	OS Average
SE4	80	57.94	72	57.43	8	62.50
HC1	188	56.71	158	56.94	30	55.53

These figures show that there is no evidence that overseas students are disadvantaged by open book examinations. Indeed, they have performed significantly better on average in SE4. Of course, a larger sample over a number of years will be needed before this can be asserted for certain.

The issue of students with conditions such as dyslexia seems no more likely to cause difficulties in an open book assessment than in a more traditional examination. Allowing the student extra time to answer the questions and considerate marking should deal with any problems.

The problems of UK students with poor skills in English is harder to deal with, and is a problem that is endemic in computing programmes. There are may cases of students in the School of Computer Studies complaining bitterly about being called upon to write essays; they see this as a job for the English students. It is inevitable that these students will suffer in open book assessments that require essay answers. However, essay answers can be required in any subject area, so this is perhaps a problem more related to the nature of the required answers than to the type of examination.

One simple step that can be taken to alleviate the problems of students with difficulties in English is simply to allow them more time to complete their answers. While they should be expected to produce answers in a reasonable time, there seems no good reason why the allowed time should be overly restrictive and should have a significant impact on their final mark.

5.3   Summary

In summary, there is no evidence from the experience presented here that the introduction of open book examinations will affect the average performance of a class of students. Furthermore, there is no serious reason to suppose that any group of students will be unfairly disadvantaged.

The group of students that are disadvantaged, of course, are the surface learners. Experience shows that, while the average marks remains similar between the types of testing, the students that score high and low marks are different. As we expect (and hope) the surface learners are at the bottom of the marks in open book examinations, and the deep learners are at the top.

6   Conclusions

This section summarises the experience and finding presented in this paper and considers the future of open book assessment in the School of Computer Studies at the University of Leeds.

6.1   The SE4 Module

Examinations in a module such as SE4 are a necessary evil. With the current expansion in student numbers, it is not possible to introduce and administer sufficient coursework to assess the module.

Open book testing has been introduced into the SE4 module. The response has been favourable from staff and, more importantly, students. The open book examination will remain in place, with a few changes. The most important of these will be to allow the students more time in the examination.

6.2   Assessing the Assessment

The question of the validity of open book assessment must be addressed. In [6] David Newble and Robert Cannon pose three questions that can be used to assess a method of assessment:

Validity - Does it measure what it is supposed to measure?

Reliability - Does it produce consistent results?

Practicability - Is it practical in terms of time and resources?

These issues will be considered in turn. From the analysis presented in section 4 it can be seen that traditional examinations successfully assess surface learning. It is also apparent that open book examinations measure something different; it is to be hoped that this is deep learning. In any form of examination there is still a significant element of "exam technique" in determining a student's mark; there is no reason to suppose that this is any less in open book examinations.

At the moment, it is impossible to say whether the SE4 examination produces consistent results. The indications from the September 1995 resit are favourable, but a larger sample will be needed before any firm conclusions may be drawn.

An open book examination is entirely practicable in most areas of computing. It takes only slightly longer to set and mark. Indeed, the students probably spend less time "cramming" and more time learning, so open book assessment may be using their time more efficiently.

As important as these issues are, one feature that has emerged from these experiences is that open book exams are more enjoyable. Some students report having enjoyed the exam (a strange concept!), and the answers were certainly more interesting to mark. The examiner certainly gains much more satisfaction from the assessment; there is genuine evidence that the students have been **taught** something (or at least have **learned** something), not that they have **memorised** something.

6.3   What do Open Book Examinations Assess?

The work presented here has shown that open book examinations successfully test deep learning. However, we must not lose sight of the fact that there is still a significant amount of examination technique involved. Indeed, there is a specific form of examination technique involved in taking an open book paper; there are new factors involved such as the ability to annotate and index notes and, crucially, the ability to decide on what material to take to the examination.

The ideal solution would be to allow the examination to be taken much more flexibly. The students might be allowed much more time, and be given access to libraries and computers if they required them. Of course, such an examination would be extremely expensive to supervise, and so is probably not practical at present.

Nevertheless, even with these shortcomings, open book assessment seems a great improvement on the traditional process.

6.4 Applicability

The results presented here are based on experiences with a particular set of students in a particular discipline, computing. There seems no reason, however, to suppose that the results would not apply equally to students in similar disciplines and notably other engineering disciplines.

It is planned to introduce open book assessment into other modules in the School of Computer Studies during the academic session 1995/96. This will provide the opportunity to gather and analyse more results to compare with those presented here. In particular, these results will be from other areas of computing and will enable the applicability of these results to be assessed further.

6.5 The Future of Assessment

In conclusion, we present some ideas on the future of assessment in computing. In [7] the authors note that "it is fair to say that every assessment method will place some students at a disadvantage to a certain extent". They go on to suggest that what is required is a range of assessment methods and hope that such an approach would lead to an assessment system that "becomes a more accurate reflection of the overall abilities ... of the students". It is already accepted that a combination of coursework and examination is used. However, it has been shown that examinations measure learning *and* examination technique and coursework is subject to plagiarism. What methods could be used instead?

One approach is that experienced by an Erasmus student from the School of Computer Studies at INSA Rouen, in France. The assessment for a programming module consisted of two one hour papers. The first involved small exercises and was "closed book". The second paper followed immediately and was open book; the exercises were longer and more detailed.

Another alternative is the use of multiple choice testing. In [8] the authors describe how multiple choice tests are effectively used in introductory programming courses. If conducted in supervised conditions, such tests are a good measure of a student's breadth of learning, although the tests can, of course, test only surface learning.

A range of assessment methods is required. The open book examination is not perfect, but it is much better than the traditional examination. It is regarded as an oddity at the moment but hopefully this will change as it is adopted more widely.

7 Acknowledgments

References

[1] Phil Race and Sally Brown. *500 Tips for Tutors*. Kogan Page, 1993.

[2] Catherine Tang. Assessment and student learning. In Graham Gibbs, editor, *Improving Student Learning*. The Oxford Centre for Staff Development, 1994.

[3] David Baume. *Developing Learner Autonomy*. Number 84 in SEDA Papers. Staff and Educational Development Association, November 1994.

[4] Phil Race. *Never Mind the Teaching Feel the Learning*. Number 80 in SEDA Papers. Staff and Educational Development Association, June 1993.

[5] R D Boyle, J Jackson, and R J Wade. Changing learning culture with electronic bulletin boards. Research Report 95.2, School of Computer Studies, University of Leeds, January 1995. Presented to *CAL-95, University of Cambridge, April 1995*. Submitted to *Computers in Education*.

[6] David Newble and Robert Cannon. *A Handbook for Teachers in Universities and Colleges*. Kogan Page, 1989.

[7] Sally Brown, Chris Rust, and Graham Gibbs. *Strategies for Diversifying Assessment in Higher Education*. The Oxford Centre for Staff Development, 1994.

[8] James Anderson, Susan Walmsley, and Shirley Williams. *Experience of Multiple Choice Testing*. Presented at the *3rd Annual Conference on the Teaching of Computing, Dublin City University, August/September 1995*.

Appendix - 1995 SE4 Examination Paper

This question paper consists of
5 printed pages, each of which
is identified by the Code Num-
ber COMP220001

University of Leeds
February 1995

Examination for the degree of BSc or BA

SE4: Commercial Systems Development

Time allowed: 2 Hours

Answer THREE questions.

Candidates are permitted to bring with them any notes (hand-written or computer-generated) that they may have. Photocopies of other material are also permitted.

Books, other than dictionaries, are not permitted.

Calculators are permitted.

You are reminded that, in normal circumstances, there are no marks available for the direct reproduction of the lecture notes.

Turn over for Question 1

Turn over

Question 1

Describe in detail how the tools provided by Ingres may be used to develop a basic prototype of a Transaction Processing System.

(12 marks)

Describe the steps involved in developing such a prototype system into a production quality system. You may assume that the production system will also use the Ingres Database Management facilities, but not necessarily the Ingres Application Development facilities.

(8 Marks)

Question 2

VIZsoft Revisited

Note: This question is based on the scenario used in the assessed coursework. Relevant extracts from the coursework specification follow.

Background

After graduating from the University of Leeds, you have been employed as a graduate trainee by VIZsoft Ltd, a well known software house. You have been through an induction week, and are about to start your first assignment.

VIZsoft have recently secured a contract to develop a system for Mail-a-Tape, a well known mail order "Music Club". This project is seen as key to the future of VIZsoft. The value of the first phase of the project is some £250k, and it is known that there is a budget of over £1m available for future work in other divisions of the mail order company.

After some preliminary analysis, it has been suggested that an approach using Fourth Generation Languages is appropriate for the task. It also seems likely that the final system will be delivered as a series of prototypes.

You are summoned into the office of the managing director if VIZsoft, Mr Bigvern. He explains the importance of the project, and tells you that a preliminary database design has been produced. The current problem, he says, is the choice of the Fourth Generation tool to be used for the project. He has a shortlist of several 4GL packages, including Ingres. He is asking the new intake of graduate trainess to undertake the evaluation of these packages for the project.

An evaluation copy of Ingres has been installed. It is a fully functioning version, running under Sun Solaris, the likely final operating system for the project.

A few weeks after you have completed your assessment of Ingres for the Mail-a-Tape project you hear that Mr Bigvern has decided to use Ingres for the project. He has also made you a member of the project team for the first phase of the project.

The project team consists of:

- A senior project manager who also works on several other projects;
- Two experienced 4GL developers;
- Two other graduate trainees;
- Yourself;

Turn over

None of the members of the project team have any previous experience of Ingres. No members of the team have ever worked together before.

The first project deliverable is due in four month's time. It is estimated that this deliverable will take roughly two months elapsed time to design and develop.

Describe in detail **four** activities that you would expect (or want) to be involved in before the design and development of the first deliverable starts. Briefly explain why you think each of these activities is important.

(4 marks each)

If it were possible for you to participate in only **two** of these activities which would you regard as the most important? Briefly explain your reasons.

(4 marks)

Turn over

Question 3

North and East Riding Digital Systems (NERDS) are a large systems house based in North Yorkshire. Their main customers are large regional public utilities such as Yorkshire Water and the Yorkshire Electricity Board. NERDS has a staff of some two hundred systems developers, and specialises in the development of large transaction processing software in COBOL.

The Systems Development Manager of NERDS, Mr Cycle, has been in his post for many years and has always favoured a traditional approach to systems development. This approach is based on a rigorous application of the Systems Development Life Cycle together with some formal structured methods.

Recently there have been problems with the development time of some projects. Customers have not been happy to wait for several months (or even years) before taking delivery of their systems or modification. Mr Cycle has decided that something must be done to improve the speed with which systems are developed.

Write a report for Mr Cycle describing how the tools and techniques discussed in this module may be of assistance to him. Mr Cycle has read many articles on these matters, so there is no need to describe the tools and techniques in any great detail; you should concentrate on suggesting which of them might be most effectively applied in this case. You should be sure, though, to analyse the basic causes of Mr Cycle's problems.

(20 marks)

(of which 2 for presentation)

Turn over

Question 4

In Commercial Systems Development the human-related problems (at both a personal and organisational level) are often more important, and harder to overcome, than the purely technical issues.

Discuss the validity of this statement in relation to **FOUR** of the following:

- Problems with the traditional Systems Development Life Cycle;
- End-User Computing;
- Systems development using a Prototyping Methodology;
- Information Centres;
- A long term systems development project where the customer site is a large distance from the base of the systems developers.

(5 marks each)

THE END

As such, open educational resources, prior learning assessment, open degree plans, and competency-based learning have gained popularity in the field of higher education. Enhancing Education Through Open Degree Programs and Prior Learning Assessment is a critical scholarly resource that examines teaching and learning materials that learners may freely use and reuse without charge. Featuring coverage on a broad range of topics such as open educational resources, prior learning assessment, and competency-based learning, this book is geared towards educations, professionals, school administrators, Open Book Assessment in Computing Degree Programmes. October 1995. Tony Jenkins. The use of "Open Book" assessment is widely seen as an approach that more effectively tests a student's understanding and knowledge of a subject area. There have been a number of attempts to introduce such assessment in the School of Computer Studies at the University of Leeds, and open book examinations are now used in several modules. This paper describes the experience of introducing open book [Show full abstract] assessment into a level 2 computing module. We consider the process of setting First and foremost, however, we hope that MSIS 2016 will serve as many schools and programs as possible in their continuous efforts to offer their students as strong, innovative, and long-lasting educational experience as possible. We encourage you to share your experiences with the MSIS 2016 model with the rest of the global IS communityâ€" please see msis2016.org to find the best ways to do so. Â MSIS-ii. Executive Summary. MSIS1 2016 provides a competency model and curriculum guidance for masterâ€™s level degree programs in information systems (IS). This joint effort by AIS and ACM builds on the foundation of four earlier graduate IS curriculum