Object-Oriented Analysis, Criterion Referencing, and Bloom

Ilona Box

School of Computing and Information Technology
University of Western Sydney
Locked Bag 1797, Penrith South DC 1797, New South Wales

i.box@uws.edu.au

Abstract

Lister and Leaney (Lister and Leaney, 2003b) presented a "coherent, explicit grading philosophy based on Bloom's taxonomy". I applied this philosophy to a second year, first semester subject in object-oriented analysis. My motivation to do so was based on a complexity of issues similar to those encountered by Lister *et al*. I report on the implementation of the philosophy, the lessons learned, as well as the fairness, equity, and feasibility of the implementation. I recommend the application of the grading philosophy as it yields a viable, repeatable, sustainable, and honest assessment strategy.

Keywords: Bloom's taxonomy, criteria referencing, software development

1. Introduction

My experience at college during my degree from 1979 to 1981 was of studying with academically committed fellow students, there was just a handful that was not. I was flummoxed to find these proportions had inversed when I began teaching in the late Nineties. I read Biggs (1999) and was relieved to find I wasn't alone; he wrote, "[The students are]…at university not out of a driving curiosity about a particular subject or a burning ambition to excel in a particular profession, but to obtain a qualification for a job…[have] a less developed background of relevant knowledge…[come] to the lecture with few questions…[and want] only to put in sufficient effort to pass". This set me on a path to do something about that situation.

The hardest issue, with which I had to come to terms, was not all my students wanted to strive for the top grade, a high distinction. I have students who are happy with any grade, so long as they pass. Sometimes, they are even disappointed if they have expended too much effort and received a mark higher than 50. Also, they work for an income a bank more than I, as a student, needed to, the average being 15 hours per week (McInnis and Hartley, 2002). I was aware that I had to improve the students learning experience, and did ((Box, 2000, 2001a, b, 2002, 2003a, b, Box and Ferguson, 2001, 2002). As well as keeping these enhancements in place, I had to come to terms with the students' diversity of ideas about what they considered as acceptable academic commitment. A lengthy discussion with Lister (2003) led me to adjusting the existing assessment strategy of the unit I coordinated, Object-Oriented Analysis (OOA).

Lister and Leaney presented a "coherent, explicit grading philosophy based on Bloom's taxonomy" (Lister and Leaney, 2003b). They devised a criterion-referenced grading scheme to address issues such as students not being able to program after introductory programming subjects (McCracken et al., 2001); "truth in sentencing" (or lack thereof) (Lister, 2001); students not achieving their true potential via norm-referenced grading, and; changes to curriculum, e.g. The Joint Task Force on Computing Curricula *et al* (2001).

I faced a similar complexity of issues, such as: students not being able to perform tasks, or exhibit an expected level of knowledge or skill in object-oriented analysis; a lack of "truth in sentencing"; stronger students not being challenged; large student numbers; multiple campuses; and the difficulties of marking and timeliness associated with the latter two issues.

After discussion with Lister (2003) where my concerns about equal opportunity, fairness, and feasibility were appeased I was motivated to implement the criterion-referenced grading scheme for a second year, first semester subject in object-oriented analysis.

I report here on the implementation of the criterion referenced grading scheme for summative assessment. I describe the lessons that were learnt and how concerns of fairness, equal opportunity, and feasibility were addressed.

2. Lister & Leaney's Grading Philosophy

Lister and Leaney's grading philosophy is to realise the potential of each student. Grading is done by assigning a grade to a student according to how that student meets explicit clear criteria, "irrespective of the resultant grade distribution" (Lister and Leaney, 2003b). The philosophy also allows students to strive for their own chosen grade.

The explicit and clear criteria are based on Bloom's Taxonomy of educational objectives (Bloom and Krathowl, 1956). The six levels of the taxonomy are grouped into three pairs and match against three passing grades, thus: knowledge and comprehension for pass grade; application and analysis for a credit or distinction grade, and; synthesis and evaluation for a high distinction. The broad criteria for each grade are expressed using verbs (examples can be found in Kemp

(1996)) that indicated the type of cognitive behaviour that was expected from a student with that particular grade. The broad criteria are also referred to as objectives or learning outcomes.

A grade is attempted by students doing tasks depending on the grade they want. For a pass grade all the assessment tasks are: laboratory exercises to have the students actively engage with the material presented at lectures; an invigilated laboratory exam, graded as pass/fail, and; a multiple-choice question (MCQ) exam with questions modelled on a style described in (Lister, 2001). For a credit or distinction a structured assignment and the pass grade assessment tasks are done to a set standard. For a high distinction a project of the students' choice, including a peer-review, and the credit or distinction assessment tasks are done to a set standard.

3. Issues and Motivation

The complexity of issues I faced: students not performing, lack of "truth in sentencing", lack of challenges for stronger students, large student numbers and multiple campuses and associated problems, was realised in a number of ways. Students were not able to perform object-oriented analysis and design tasks, or exhibit an expected level of knowledge or skill. This was most tellingly revealed by the lecturer of the capstone project subject,. he said, "I start classes with the assumption that the students know nothing." This was based on his experience of lowering his expectations of students' abilities doing the capstone project subject.

Lister *et al* (Lister and Leaney, 2003a) cite evidence (McCracken et al., 2001) that there is a problem with the way programming is taught. Though similar evidence in the object-oriented software development discipline was not found, anecdotal evidence (Almgren, 2002) and personal experience suggests that the same problem exists with the teaching in this discipline.

The lack of "truth in sentencing" (Lister, 2001) is when the marks and/or marking criteria of assignments and exams are not closely matched to stated objectives of the subject or the tasks themselves. This lack combined with students' compartmentalizing of subjects was evidenced in a number of ways. For instance, students acted as if they had not done a university standard assessment task before, even though they were in their second year, or a subject in the principles of professional communication. Again, anecdotal evidence (Editor, 2003) suggests that this issue is real and common.

The strongest student not being challenged was another issue I faced. There are differences between weaker and stronger students. The weaker students may not be poor academically in all subjects; they may just be poor in this subject due to their lack of talent or interest in the content. The stronger students are usually interested and quite good academically; they sometimes have a talent for the subject content. For an academic who is passionate about ones specialty it can be difficult to imagine that others are not; for those that are not interested, it is merely a question of exposing them to the "beauty" of the specialty to become so. This is not reality. Not all students are interested in all the core subjects in a degree. Hopefully, they are able to choose electives in which they are interested, if not talented. For the weaker students we need to recognise their marginal interest; stronger students need challenges to nurture their talent and spur their interest.

Large student numbers is an ongoing issue. Enrolment numbers in the hundreds is common; OOA starting enrolment was over 500. The subject, OOOA, is taught on three campuses. It takes approximately an hour to drive from one campus to another. The logistics of marking, timeliness of feedback, consistency of marking, moderating marks, equity across campuses, evenness of delivery, quality and availability of sessional staff, and methods of access to subject materials are all impacted by large student numbers and multiple campuses.

I had addressed some of these issues by using formative assessment that engaged students in learning (Box, 2002, 2003a, b) and providing a text appropriate for the students (Box and Ferguson, 2002). I felt more could be done. I was motivated to explicitly address the issues by implementing an honest, realistic, achievable, fair, and pedagogically sound, summative assessment strategy. The issues faced by Lister *et al* were similar to mine. They had addressed their issues by adopting a grading philosophy that provided a suitable summative assessment strategy. I chose to implement their strategy.

4. Implementation

Object-Oriented Analysis (OOA) is a second year subject with an initial enrolment of about 500 students; in Autumn 2003 just over 300 sat the final exam. (The attrition was largely due to incorrect enrolments, in particular not meeting the prerequisite.) OOA is taught on three campuses. One subject or unit coordinator (developer and provider of subject materials and management), two lecturers (one of whom was sessional staff), and seven tutors (all sessional staff) were the staff used to deliver OOA in Autumn 2003.

The implementation of the Lister and Leaney grading philosophy required changes to the unit objectives, exam structure, continuous summative assessment tasks, and determination of the final grade.

4.1 Unit Objectives

The existing unit objectives had been written with Bloom's taxonomy for educational objectives in mind. However, changes were required to identify and group which objectives applied to which grade and by carefully choosing some words (shown in italics in Figure 1) so that the objectives matched Bloom's taxonomy and the corresponding grade. The resulting unit objectives (Figure 1) were improved as they were more closely aligned to the assessment tasks and embodied the levels of understanding in advance (Biggs, 1999).

At the conclusion of Object-Oriented Analysis students who have achieved: - At least a Pass grade will be able to:

a) Exhibit knowledge of the purpose and meaning of a development lifecycle and the phases in an object-oriented software development lifecycle by accurately *identifying* these.

b) Accurately *recognize* most of the object-oriented information system modelling and development techniques in the early phases of system development for a given case study and possibly other smaller scenarios.

c) Accurately *recognize* the appropriateness and correctness of artefacts in the early phases of software development for a given case study and possibly other smaller scenarios.

d) Use a CASE tool to *reproduce* artefacts applicable to the early phases of an object-oriented software development process.

e) Use selected features of an object-oriented integrated development environment to *reproduce* artefacts applicable to the early phases of an object-oriented software development process.

f) Use selected features of a word processor to *reproduce* artefacts applicable to the early phases of an object-oriented software development process.

A Credit or Distinction grade shall satisfy the objectives for a Pass and will be able to:

g) *Apply* knowledge of object-oriented software development by writing portions of successful, accurate, and workable deliverables for the early phases of the development process based on their *analysis* of a small case study.

h) Demonstrate a beginning developer's level of *questioning* skills by *practicing* the skills when participating in class and gathering requirements from the client for the case study and appropriately including their findings in the deliverables.

i) Accurately *interpret* verbal and written instructions about the performance of software development tasks and associated administrative duties.

A High Distinction grade shall satisfy the objectives for a Credit or Distinction grade and will be able to:

j) *Prepare* a project proposal for a small software system.

k) *Synthesize* knowledge of object-oriented software development by *creating* successful, accurate, and workable deliverables for the early phases of the development process for the proposed project as well-documented artefacts.

l) *Evaluate* the quality of artefacts produced from the early phases of an object-oriented software development process.

Figure 1. Objectives of Object-Oriented Analysis Autumn 2003 (Words in italics relate to Bloom's taxonomy.)

4.2 Exam Structures

Lister *et al*. use two exams, a laboratory exam, which I called a practical exam, and a MCQ final exam.

4.2.1 Practical Exam

The pass/fail practical exam replaced the, often gratuitous, small percentage points for continuous tutorial assessment. The weekly tutorials were used for formative assessment as described in (Box, 2002, 2003a, b) . Each student was encouraged to develop an electronic portfolio of the activities. (The portfolio could later be used as a demonstration of a student's learning if they wish to contest their grade.) The preparation and in-clas s activities prepared the students for the practical exam.

The practical exam addressed objectives d, e, and f (Figure 1). Ten different possible questions were set and made available online, as were all unit material, prior to the practical exam. Every student on every campus was able to prepare for the practical exam based on the same material. This was a fair and equitable way of addressing the issue of multiple campuses (and access to unit materials). Each student was set one of the 10 questions, at random, at the time of the exam. The tutor and students were given the marking criteria before the exam, which can be broadly stated as: the student's reproduced artefact must exactly match the provided artefact. The tutor marked the exam in the tutorial session and made the mark immediately available to the student. Any student who failed the practical exam was able to re-sit the following week.

Students were encouraged to practice all ten questions. Even though they may not be tested on all three objectives the practical exam addressed, as the questions were randomly assigned at the time of the exam, they should have been able to meet all three objectives.

As the questions had been made available beforehand plagiarism was not an issue. The exam was held in the usual tutorial sessions, in the last two weeks of the semester.

The practical exams took just over three hours to write; it was an edited version of the tutorial materials, which took approximately 65 hours to prepare.

In keeping with Lister *et al* (Lister and Leaney, 2003a), approximately 90 percent of the class passed the practical exam at their first attempt, almost everybody passed after their second attempt.

4.2.2 MCQ Final Exam

Previous final exams included a MCQ section, short answer questions, and a case study. This was an attempt to test the students' competency in all the objectives in a controlled environment. This is in itself a poor assessment design. The marking of the non-MCQ portions was always problematic because of the large number of papers, the fatigue of the marker, consistency of the marking, the moderation of the marks, and the logistics of getting the papers to and from markers, if more than one marker was involved. When I took on the

marking is a solo effort it often involved long hours and weekends, placing unrealistic demands on myself, consistently beyond what most would consider a reasonable workload.

Another issue with the non-MCQ portion of the final exam was the lessening of the "truth in sentencing". Grace points were given for written answers to the previous final exam. Marking approximately 2000 pages of (often poor) handwriting, to a strict deadline, proved an arduous task and to re-mark questions was not viable or sustainable. Accumulating grace points contributed to the lack of “truth in sentencing”. It was difficult to share the marking as this increased the risk of losing papers or required several markers to be in the same place for many hours, and it was difficult to find good markers. Under the circumstances, to err in the favour of the students by giving marks for their writing anything, even a word or two, related to the question and desired answer was more likely than giving a zero (especially when there is pressure from various levels in the institution not to fail to many students). To remove these problems with the non-MCQ component of an exam is easy, move to a (preferably) machine-marked totally MCQ exam.

The move to an entirely multiple-choice question exam is perhaps the most contentious issue with Lister and Leaney's assessment strategy. Lister’s presentations, though often received favourably, usually are followed with objections about the MCQ exam. Yet the reality is it is not sustainable practice to have workload estimates of 150 to 200 hours of exam marking at the end of every semester. Nor is it necessary for all objectives to be tested at all levels of competency in one assessment task. The MCQ exam tests objectives at the knowledge and comprehension levels of Bloom's taxonomy. And the MCQ exam only tests some, not all the objectives. The MCQ final exam assessed objectives a, b, and c (Figure 1).

An exam based entirely on well-written MCQs requires a considerable amount of time spent formulating a reliable (KR20 of at least 0.8) and valid question bank. Lister spent over a week (Lister, 2000) writing an exam of 40 MCQs. I spent 33 hours writing 60 MCQs, which yielded a KR20 of 0.78 for the Autumn 2003 exam.

The pre-set standard for passing the MCQ exam was 70%. Lister *et al* also set such a high pass mark. One reason is the calculation of a passing mark: if the pass mark for a pass grade is 50 out of 100 and the MCQ exam is worth 70 of the 100 marks then 70% of 70 is 49, which is near enough to 50 to be the required passing standard. Another reason is the level being tested is knowledge and comprehension. I, as to others, expect students to have a high level of knowledge and comprehension even though other levels in Bloom's taxonomy may not be met at a similar level.

The results of the MCQ exam need careful consideration, which takes considerably less time than marking written answers. Using the raw score as the starting point, then adjusting for questions with two correct answers, just over 17% of students achieved 70%. The standard was lowered, according to the determination of final grade stipulated in the unit outline discussed later, to 53%. This was disappointing. I may be able to improve this figure by providing examples of the MCQs to students during the semester. They will then have had practice at doing these sorts of questions.

One distinct advantage of a MCQ exam is that it can be machine marked. Then, not only are the results of the students provided, but also statistics such as the reliability of the exam and an item analysis are reported. Though the result processing timeline did not allow the complete and in-depth consideration of these statistics to influence the results of Autumn 2003, further work is being done, using the multiple-choice question report, to develop a process to do this quickly. So far, it has been found that using the statistics in the computer marking report can increase reliability and validity (early indications are that the reliability could be marginally increased and the pass mark raised by as much as 11 percent). Also, when reworking the questions for future exams the item analysis can be used to help discover the elements of good questions. If identifiable these elements can be transferred to other questions.

4.3 Continuous Summative Assessment

As stated earlier a structure for formative assessment to engage the students in active learning was already in place. It was felt that the existing continuous, summative assessments required changing.

The existing continuous, summative assessment was standard across most units in the school, i.e. two or three assignments, the first due at week 5 or 6, the second at week 9 or 10 and the third, if set, due at week 12 or 13 of a 13 week semester. All the assignments were compulsory. This meant the weaker students submitted work that was difficult to comment on as there was often so much that needed comments that they could amount to more than the submitted assignment. Also the weaker students tended to gain grace points for effort rather than quality and the same problems with marking the written component of the final exam exist with marking these assignments. The strongest students were not challenged by these assignments.

The implementation of the Lister and Leaney grading philosophy meant that continuous summative assessment tasks were no longer compulsory. This meant the weaker students, especially those not particularly interested in this subject, could opt for assessment by the two exams only.

The two continuous, summative assessment tasks were due on the same date. This meant that only the stronger students, with appropriate time management, were likely to complete the two tasks. The two tasks were a structured assignment and an individual project.

4.3.1 Structured Assignment

The structured assignments assessed objectives g, h, and i (Figure 1). The structured assignment was similar to earlier assignments set in previous semesters. The main

differences that needed to be made so that the assignment was similar in type to that used by Lister *et al* were reducing the amount of work, not working in groups, and explicitly limiting the scope to six use cases.

For earlier assignments, students tended to produce large documents, approximately 20 pages each. Consequently, to make marking feasible and the work achievable by the students, assignments were done in groups. This often resulted in students relying on others to do the work, disgruntled committed students, and unsatisfying learning experiences all round. Limiting the amount of required work, removing the need to work in groups, and the compulsion that assignments must be done by all, meant the marking became more manageable and only those students interested in achieving higher grades deciding to do the work.

Of the 134 assignments submitted, 50 scored eight or more out of 14, 28 scored six, and 56 scored three out of 14. Again a disappointing result that I feel could be improved by providing more examples of what is expected.

The structured assignment details were based on previous assignments. Approximately 3.5 hours were spent preparing the details of the structured assignment for OOA in Autumn 2003. Grouped comments were provided as feedback.

4.3.2 Individual Project

The individual project assessed objectives j, k, and l (Figure 1). The project was divided into three parts: a two-page project proposal, a deliverable from the first two phases of a five-phase software development life cycle, and an evaluation of a deliverable.

The characteristics of the individual project were similar to Lister and Leaney's (Lister and Leaney, 2003b). The students were asked to develop a system of their own choice and limited tasks were set for completion.

About 150 completed the proposal, 27 students managed to complete up to part two, only 17 students managed to complete all three parts of the project. Of the 17 completed projects, 13 were of an acceptable standard. Again a better performance would have been preferred and this is likely to be achieved by providing more examples for the students of projects and assisting them with their time management.

The individual project details were based on previous assignments. Approximately 1 hour was spent preparing the details of the individual project for OOA in Autumn 2003. Another hour was spent writing group comments as feedback to the proposals. Each individual project part two and three received personalized comments.

The short amount of time needed to prepare the assignments was due to the ongoing evolution of the assignments since 1997. A template I had written, now also in used by others, is easily edited to suit. The details of content specific to the project or assignment will all that needed adding and editing.

4.4 Determining Final Grades

I am required to stipulate how the final grade is determined in the unit outline. According to guidelines set by my institution's Academic Senate, an acceptable portion of "F" grades is 20% of enrolled students. Also, I had a prior occurrence of 58% of students failing. From this event, I learnt that at least 60% of students must pass, or the grades were likely to be changed. The Academic Senate ruling, the previous event, and as the assessment structure, and to some extent the assessment tasks, were new to the students accommodations were made to ensure the percentage of students passing was at least 60%.

Only the pass grade assessment tasks were compulsory. I had used the tutorial preparation and in-class activities, to equate to Lister and Leaney's laboratory exercises, as ungraded formative assessment. The practical exam was marked as pass/fail. This left the MCQ final exam mark as the one that could be manipulated to achieve an acceptable portion of passing students and therefore determine final grades. As stated earlier the pass mark for the MCQ final exam was lowered from 70% to 53%.

Unlike the student cohort taught by Lister and Leaney, which is drawn from the top 12% of high school students, my students are drawn from the top 35%, most of who fall between this 12% and 35%. This means I have more students that are not as experienced at achieving high academic standards. Class time was spent explaining the assessment structure and this was reinforced in a number of places in the unit materials. However, I felt that more could be done to help the students understand the links between the objectives, assessment tasks, and the determination of the final grade.

5. RECOMMENDATIONS

My experience of implementing the Lister and Leaney grading philosophy provided me with a number of lessons that I have set out here as recommendations. I consider that these recommendations, though not exhaustive, are applicable to others choosing to adopt the Lister and Leaney grading philosophy.

1) Allow more time than one might expect is required to prepare high-quality assessment tasks. Though certainly not empirical evidence, the assignment and project, on which I spent the least amount of time, seemed to be more poorly interpreted by the students than other subject materials. This also needs to be considered with my experience that the students are neither sets standards nor do markers adhere to the standard when set. Consequently, the interpretation of what was required is in part a reflection of the students experience with previous and similar tasks.

2) Provide examples of all assessment tasks. My students had not encountered such a rigorous or demanding assessment strategy and therefore were unfamiliar with the quality and nature of the tasks. For the formative and continuous summative assessment provide examples of completed assignments and

projects similar to the ones that are set as graded or ungraded tasks. For the exams provide example questions during the semester, preferably each week, for practice. Hence, the example questions can be built up over time. The seen MCQs can also be used in the exam to build the students confidence as they encounter something familiar while doing the exam. It is recommended that for good students at least three worked examples of the tasks or style of question be provided and for average students at least six examples (Donald, 2003). This might be considered arduous. However, it provides a more manageable workload by spreading the work out and allowing an accumulation of quality materials. A planned approach, over more than one semester, so the development and improvement is incremental, is a sound strategy, resulting in a realistic workload and achievable outcomes.

3) Make well informed, planned, and well timed decisions about the objectives:

a) It is necessary to decide either to sacrifice some content to deepen learning or to sacrifice depth of learning for breadth of coverage. A liveable medium needs to be found, expressed in the objectives, and reflected in the size and complexity of all assessment tasks. I strongly recommend reading about this aspect, for instance, Biggs(1999), before taking action.

b) Write really good objectives that adhere to the recommendations for well-written objectives. Again read about this, for instance, (Bloom *et al.*, 1971), before taking action.

c) Work carefully at grouping the objectives into the appropriate grade. This is a key activity to documenting the real, achievable, and honest objectives. I felt I had done this reasonably well.

d) Carefully choose the words that reflect the level of cognitive behaviour and include these in the objectives. Writing objectives requires practice. I found that the objectives could be greatly improved by reviewing the literature about writing objectives and referring to good examples. Be sure that the right word is used, in the right place, and in the right way.

e) Include a first objective, in the pass grade objectives, that lists the terms and vocabulary one thinks a student should have by the end of the subject. This is useful, as well as good practice, as it makes it clear to the students that if they cannot understand the remaining objectives because they include vocabulary particular to the subject, then the students should understand the objectives as the material is covered. This list also explains to the students that if they do not understand the terms used in the rest of the objectives that they should come to understand the objectives during the subject. This list of terms is also fodder for some of the MCQs in the final exam.

4) Take the time to explain to the students the grading philosophy including such things as the meaning of criteria referencing and the links between the objectives, assessment tasks, and marking criteria. Provide examples were doing this.

5) Allow more time than perhaps had been previously allocated to working up good practical exam questions and MCQs. The Lister and Leaney grading philosophy necessitates effort to develop quality, reliable, and valid exam questions.

6) Well-written MCQs are likely to be unfamiliar to the students. For instance, one student, after spending six hours with the OOA exam writer reviewing the exam, commented that she doubted if she was a 70% plus student, indicated by previous grades, as she recognized she had never sat an exam of quality MCQs. She felt that she would have benefited from practice and early exposure to the style of questions, as well as using the provided exam preparation material.

7) Be humble enough to acknowledge that you don't know all about or need practice at learning what is and how to write good MCQs. Take the time to read papers and books about MCQs and find examples to compare. Do not be too accepting of any one version of what is a good multiple-choice question. It was found that much of what is reported about good multiple-choice questions has not been supported by research or measured.

8) Make any preliminary work on continuous summative tasks, such as preparing a project proposal, formative and voluntary. This should distinguish the stronger students from the rest of the class early, so that attention can be directed there, and ease the marking load. (The effort perhaps better directed at upfront preparation of subject material, assessment tasks, and assessment examples.)

9) Take the time to understand and use the statistics provided by the MCQ machine marker. Use the statistics produced from computer marking programs of MCQs to improve the reliability and validity of the exam before determining the final grade. (This is discussed in depth in a forthcoming paper by the author.)

10) The structured assignment could be subject to plagiarism. This might be overcome by “designing out opportunities for plagiarism”, such as “individualizing the task” (Carroll and Appleton, 2001). For instance, this might be achieved by letting the students choose the system and setting strict limits on the content such as only six use cases, a maximum of nine classes, and only five attributes maximum per class. However, this impinges on the higher cognitive behaviours demanded in the individual project. Another alternative may be to take a very large system and slice it into portions, for example one or two use cases per student. This does require more management and administration.

12) Employ sessional staff who are trained and who you can trust to work in harmony with the grading philosophy and other innovations you may use. I took

care to develop lesson plans and made them available to staff and students so all were informed of the format and content of lessons. One student wrote "[I would like less of] The lesson plan. It was confusing as we did not follow it.". This suggests one or more sessional staff undermined the equity of delivery across campuses by deviating noticeably from the lesson plan.

13) Allow two sittings of the final exam. This could be difficult to arrange through the formal institution examination process. However, it is worth giving the students the opportunity to sit the exam twice. The students are given the opportunity to pass in a more genuine manner than having to scale their score.

14) Use your imagination. Break away from stale practices. Try new things. Talk to like-minded people to get some direction for new paths and support, for instance, e-mail or phone me when stumped for multiple-choice questions.

6. CONCLUSION

I recommend the application of the Lister and Leaney grading philosophy. I also think it is important for others to report on their experience with the assessment strategy so we can support each other.

I would act on the 14 recommendations that have been made as a result of my first implementation of the Lister and Leaney grading philosophy, especially allowing time to develop quality assessment tasks and worked examples.

I did not expect to have unbridled success with my first implementation of the grading philosophy. I see that part of teaching is the evolution of good practice. Therefore, I will continue to reflect on my experience with the assessment approach.

I believe there is great benefit to the Lister and Leaney grading philosophy. Teachers with issues similar to those mentioned can consider the grading philosophy, and made a thoughtful implementation of it, as it will yield a viable, sustainable, honest, and repeatable assessment strategy.

7. REFERENCES

ALMGREN, P. (2002): Personal communication: e-mail. with BOX, I., 15 November

BIGGS, J. (1999): *Teaching for quality learning at university*. Buckingham, England, Society for Research into Higher Education and Open University.

BLOOM, B. S., HASTINGS, J. T. and MADAUS, G. F. (1971): *Handbook on formative and summative evaluation of student learning*. New York, McGraw-Hill.

BLOOM, B. S. and KRATHOWL, D. R. (1956): *Taxonomy of educational objectives*. New York, NY, USA, McKay & Co.

BOX, I. (2000): Engaging students in information systems development - a detailed account of first encounters, *Fourth Australasian Computing Education Conference*, Melbourne, Vic, Australia, Vol.:253, ACM.

BOX, I. (2001a): Engaging students in learning: It's about what they do, *Teaching and Learning Conference 2001*, Northern Territory University, Darwin, Vol.:53-60, Teaching and Learning Development Office, Northern Territory University, Darwin, NT, Australia.

BOX, I. (2001b): *Integrating first year subjects*. The ICT-Ed Project: Learning Outcomes and Curriculum Development in Major Disciplines in Information and Communication Technology Available at http://cerg.csse.monash.edu.au/icted/briefaccounts.html#uws2. Accessed on 15 October 2003.

BOX, I. (2002): Applying educational research to improve teaching and learning in information systems, *Informing Science and IT Education Conference*, Cork, Ireland, Vol.,

BOX, I. (2003a): Assessing the assessment: An empirical study of an information systems development subject, *Australasian Computing Education*, Adelaide, SA, Australia, Vol.,

BOX, I. (2003b): Using formative assessment and the feedback process as an approach to the learning of software development review and quality assurance skills, *HERDSA*, Christchurch, NZ, Vol.,

BOX, I. and FERGUSON, J. (2001): Improving software development: The prescriptive simplified method, *Australian Conference on Information Systems*, Coffs Harbour, NSW, Australia, Vol.,

BOX, I. and FERGUSON, J. (2002): *Object oriented software development: Step by step*. Sydney, Pearson Education.

CARROLL, J. and APPLETON, J. (2001): *Plagiarism: A good practice guide.* Oxford Brookes University Available at http://online.northumbria.ac.uk/faculties/art/information_studies/Imri/Jiscpas/docs/brookes/brookes.pdf. Accessed on 12 September 2003.

DONALD, J. (2003): Keynote address, *HERDSA 2003*, Christchurch, New Zealand, Vol., Higher Education Research and Development Society of Australasia.

EDITOR (2003): *Keeping degrees at a high standard*, The Australian. 18 July. Editorial.

KEMP, K. (1996): *Ncgia giscc learning outcome: Writing learning outcomes for the core curriculum.* Available at

http://www.ncgia.ucsb.edu/education/curricula/giscc/units/format/outcomes.html. Accessed on 12 September 2003.

LISTER, R. (2000): On blooming first-year programming, and its blooming assessment, *ACE*, Melbourne, Australia, Vol., ACM.

LISTER, R. (2001): Objectives and objective assessment in cs1, *SIGCSE*, Charlotte, NC, USA, Vol., ACM.

LISTER, R. (2003): Personal communication: meeting. with BOX, I., 26th January

LISTER, R. and LEANEY, J. (2003a): First year programming: Let all the flowers bloom, *5th Australasian Computer Education Conference*, Adelaide, SA, Australia, Vol. Conferences in Research and Practice in Information Technology, Vol 20,

LISTER, R. and LEANEY, J. (2003b): Introductory programming, criterion-referencing, and bloom, *SIGCSE'03*, Reno, Nevada, USA, Vol., ACM.

MCCRACKEN, M., ALMSTRUM, V., DIAZ, D., GUZDIAL, M., HAGAN, D., KOLIKANT, Y. B.-D., LAXER, C., THOMAS, L., UTTING, I. and WILUSZ, T. (2001): Iticse 2001 working group reports: A multi-national, multi-institutional study of assessment of programming skills of first-year cs students. *ACM SIGCSE Bulletin,* **33**(4):125-180.

MCINNIS, C. and HARTLEY, R. (2002): Department of Education, Science and Training, Canberra.

THE JOINT TASK FORCE ON COMPUTING CURRICULA, IEEE COMPUTER SOCIETY and ASSOCIATION FOR COMPUTING MACHINERY (2001): Computing curricula 2001: Computer science. *ACM Journal of Educational Resources in Computing,* **1**(3).