

Assessing the Assessment: an Empirical Study of an Information Systems Development Subject

Ilona Box

University of Western Sydney, School of Computing & IT

i . box@uws . edu . au

Abstract

Universities need to respond to a change in the mix of student profiles from a majority of the traditional academically committed student to a majority of the students who seek a qualification for a job. The latter type of student lacks study skills such as an ability to structure his learning experience and to engage actively and continuously with the material to be learnt. This study reports on the application of one formative assessment strategy in response to the change in the mix of student profiles. The aims were to encourage students to engage in deep learning; increase a student's responsibility for learning by providing a structure for the learning effort; and encouraging continuous study. The results indicate that the aims were to some extent achieved. Consideration of formative assessment, alternatives that might improve our response to the change, and areas for further research are identified.

Keywords: Information systems teaching and education, formative assessment and evaluation, assessment validity and reliability, course objectives.

1. Introduction

This paper is about improving learning experiences by structuring the learning environment with the use of formative assessment. Mass education has meant a change in the type of student at university (Biggs,1999). The students are, as described by Biggs, "...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 study is in response to the change in the student profile. From a student perspective, the goals of the study were to engage students in deep learning, rather than surface or strategic learning; increase the responsibility taken by each student in participating in their own achievement of the subject objectives; and encouraging the development of more constant study practices that emulate a work environment.

Research of educational assessment, though regarded as uneven and weak (Black,2000) presents a consensus that subject objectives are achieved by the use of reliable and valid assessment (e.g. Biggs,1999, Boyd,2001, Chase,1999, Crooks,1994, Kanjee,2000). Using such assessment (and responsive teaching practices) will therefore yield positive learning outcomes for the students. The goals of this study were to be achieved by using a rigorous structure of formative assessments.

An earlier paper (Box,2002) presented a broader report on the study. This paper reviews literature about formative assessment and the influence this had on the assessment strategy; critiques the application of one type of formative assessment strategy to the teaching of object-oriented software development; describes the subject, its objectives, assessment and the outcomes; relates the outcomes to the aims of the study and the literature; and the direction of further research is presented.

2. What is Formative Assessment?

Generally, assessment can be described as either summative or formative. Summative assessment takes place at the end of a unit of study and "sums up" the current level of knowledge and skill a student will have of that unit of study.

Formative evaluation was first used by Scriven (1973) as the term to describe the evaluation of curriculum while it is still under development. Formative evaluation has since been widely adopted as the term to describe "systematic evaluation in the process of curriculum construction, teaching, and learning for the purpose of improving any of these three processes" (Bloom, *et al.*,1971). Bloom *et al.* went further by describing the use of formative evaluation to improve the teaching and learning processes. Formative evaluation, in the Australian context, is the same as formative assessment.

Formative assessment is used to determine whether a learner has achieved an adequate level of skill or mastery of some subject content before the opportunity to learn that subject content has passed. This requires some judgment to be made of the student's attempt at the assessment task. Bloom *et al.* note that there is a negative aspect associated with evaluation and the "users" making such a judgment might reduce this. This supports the argument by Rolf and McPherson (Rolf and McPherson,1995) that formative assessment should be non-judgmental and non-threatening. And leads us to Bligh's (2000) observation that a series of formative assessments, where the judgment is made by the

teacher, is "a multiple set of summative assessments". However, others (Black and William,1998, Cowie and Bell,1999, Torrance,2000, Bloom, *et al.*,1971) consider the assessment task to be formative when the judgment of the student's work is made by the teacher.

Scriven states, "the evaluation feedback loop stays within the developmental agency (its consultants), and serves to improve the product". To establish what would constitute formative assessment one needs to decide what is the feedback loop, what is the developmental agency, and what is the product. By confining the possible developmental agencies to the teacher and student or just the student and the primary product to student learning, I propose a number of feedback loops (Figure 1). The first feedback loop of judicious teacher/responsive student (Figure 1a) shows that the teacher plays a role of setting the task and criteria against which the task may be judged. The student responds with an attempt at the task, using the criteria as a

guide for adequacy. The teacher grades and provides comments on the attempt at the task. The student responds with another attempt at the task. This loop can continue until the student demonstrates an adequate mastery of the task or as long as time permits (which may mean in a group of students that some do not achieve an adequate mastery of the task). The judicious teacher/responsive student feedback loop leaves the onus of judgment on the person with the most skill and subject expert knowledge. Some of the likely strengths of the judicious teacher/responsive student loop are: the judgments are more likely to be correct; the time to achieve adequate mastery, for the majority of the class, is shortened; the student can concentrate on responding to comments and doing the task. Some of the likely weaknesses of the judicious teacher/responsive student loop are: the student is not participating in making judgments and is less likely to acquire this skill; the workload of the teacher might increase and

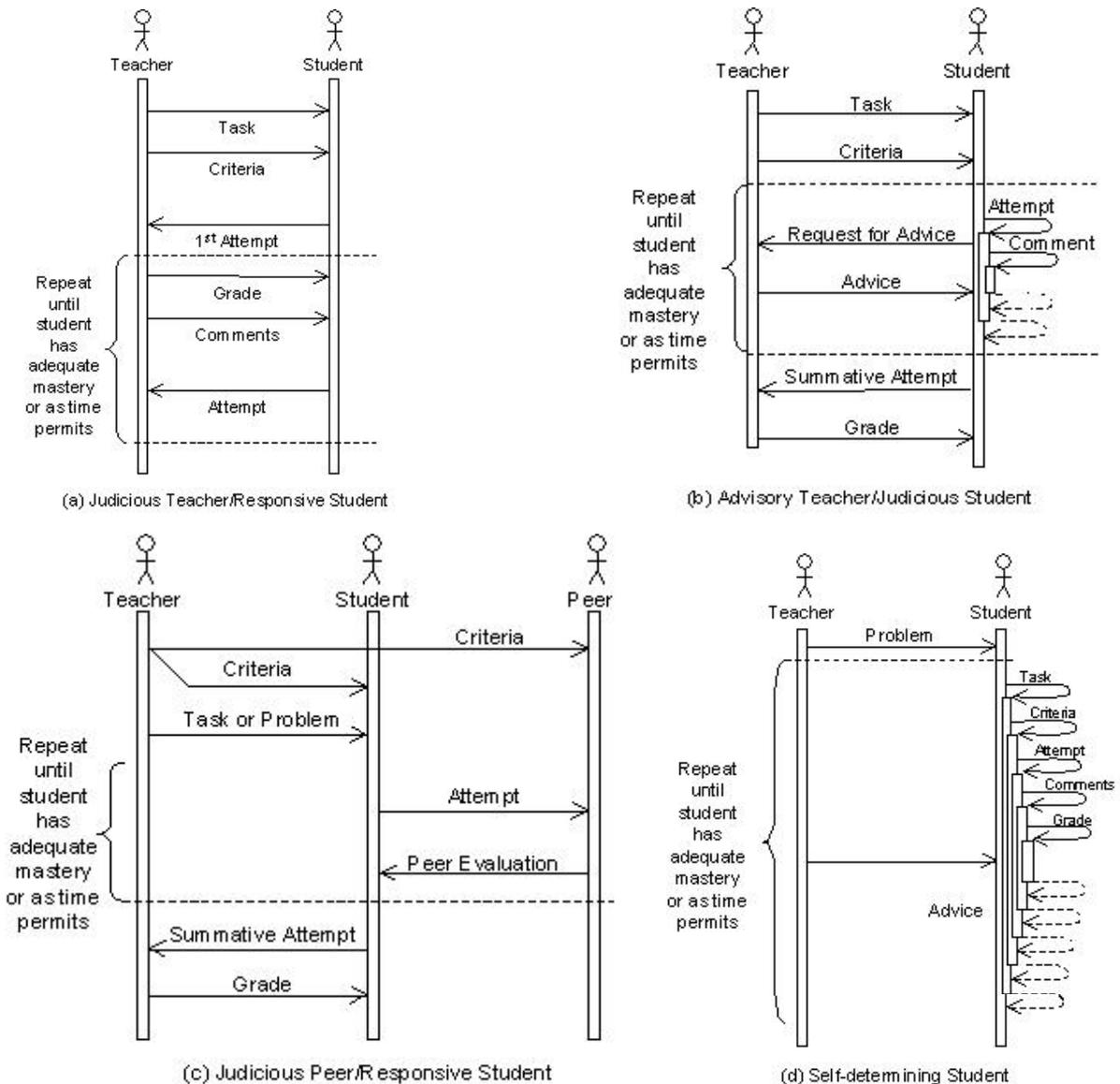


Figure 1: Assessment feedback loops.

has become more time critical.

The second feedback loop, advisory teacher/judicious student (Figure 1b), has the student in a more active role of making judgments about work set by the teacher. The teacher plays a role of setting the task and criteria against which the task may be judged. The student attempts the task and, using the criteria, comments on the attempt at the task. The student can attempt and judge the task as often as they like, or as time permits. The teacher can advise the student as much and as often as the student wants. The teacher's advice is directed more at the judgments the student is making, that is to say the student's interpretation of the criteria and how it applies to their attempt at the task, than the attempt itself. The loop would end when the task is treated as a summative assessment and the teacher grades it. Some of the likely strengths of the advisory teacher/judicious student feedback loop are: the student is participating in making judgments and is more likely to acquire this skill; the teacher may be seen by the students as more approachable and less threatening; the students can work at their own pace. Some of the likely weaknesses of the advisory teacher/judicious student loop are: the students may initially flounder at making judgments and develop a negative opinion about studying the unit; the students don't have the subject expert knowledge to make accurate judgments; the teachers may be perceived to be shirking their "traditional" responsibilities of teaching students how to do better; the coverage of subject material may be significantly reduced.

The third feedback loop, judicious peer/responsive student (Figure 1c), has the student's peer (or peers) commenting on and grading the student's work. The teacher's role may be to initiate the loop by setting a task or problem, and providing the criteria to the student and peer and/or input to the loop as determined by the students. Some of the likely strengths of the judicious peer/responsive student loop are: cooperative rather than competitive learning will be encouraged; the notion of peer review such as walkthroughs that take place in a work environment will be encouraged; the importance of peer-assessment in formative assessment will be accommodated (Black 2000). Some of the likely weaknesses of the judicious peer/responsive student are: the lack of experience of the judicious peer may lead to animosity between students; the judgments mislead the student as to the quality of their work, usually by over favorable judgments (Scriven, 1991).

The fourth feedback loop, self-determining student (Figure 1d), has the student commenting on and grading their own work. The teacher's role may be to initiate the loop by setting a problem to be solved and provide input to the loop as determined by the student. Some of the likely strengths of the self-determining student loop are: successful students will become independent learners; the teacher's role becomes explicitly one of a facilitator to learning. Some of the likely weaknesses of the self-determining student are: unsuccessful students may become hostile; teachers may feel they are too far removed from a comfortable teacher/student relationship.

The developmental agency in the first loop is seen to be equally the teacher's and students' responsibility. The responsibility of development shifts more onto the student in the second loop. In the third loop the student and peers are the developmental agency, and in the fourth loop the student is the developmental agency.

The product of all loops in the context of this study is a student who demonstrates good object-oriented software development skills and knowledge. The mix of being able to do software development tasks and assure the quality of those tasks is suspected to be more even in the third and fourth loops than in the second and more even in the second loop than the first.

Testing each feedback loop in a given situation can discover the merit of using one over another. It is through this testing the proposed likely strengths and weaknesses may be found to be true.

For this study we chose the judicious teacher/responsive student feedback loop, hence the developmental agency is the teachers and students and the product was students who demonstrated good object-oriented software development skills and knowledge. It was hoped that this feedback loop would overcome the difficulties students might experience trying to determine if their learning was satisfactory (Bloom, *et al.*, 1971). The feedback loop between teacher and student required the judgment of each assessment task to be made by the teacher, in keeping with Black et al and others mentioned above. The judgment was expressed in terms of mastery of material. It was suggested (Bloom, *et al.*, 1971) that the mastery could be indicated by a zero for inadequate, a tick for adequate, and a plus for above adequate. However, the students were more familiar with pass and fail indicators, so these were chosen. Scriven and the many authors on formative assessment that followed agree that the formative role of assessment is one of ongoing improvement. Along with an indicator of mastery, sufficient feedback needs to be provided to the students to direct them to improve their learning; just as at the end of a walkthrough a developer receives a list of errors that require correction.

The developmental agency was the teachers and students as they were the people in feedback loop and the assessment was to stay within that loop. The teachers and students were the agents who would suggest ongoing improvements and participate in the development of the product. The teacher is primarily responsible for improvements to the delivery of the material. The student is primarily responsible for the learning of the material. However, these responsibilities can be shared, at least to some small degree, by the teacher and student exchanging ideas and reflecting upon their experiences.

Scriven (1973) describes the primary, secondary, and tertiary effects of evaluation. The primary effect is that which is evident directly in the students. The secondary effect is that which is evident in those directly in contact with the material other than the students, such as the teachers. The tertiary effect is evident in the people other than those di-

rectly in contact with the material, such as teachers of other courses or subjects, and administrators.

The primary and secondary effects could be intentionally gained by identifying the final products appropriately. For this study, the primary product was a student who demonstrated good object-oriented software development skills and knowledge, the emphasis being on their ability to do the task. A secondary product was a design of a structured learning experience targeted at meeting the objectives of the subject.

Formative assessment can also be used to enforce a time schedule for each unit of study (Bloom, *et al.*,1971). When a student has an expectation that performance is required they are more likely to adequately prepare. Frequent, small, formative assessment tasks should reduce the risk of students postponing their study until just before a major or final summative assessment. This can be likened to the use of walkthroughs during system development. A developer will quite wisely ask for a formative review of their work during its development so that the developer can check their "learning" about the system. It is for these reasons that multiple formative assessment tasks were used to emulate a work environment where constant effort and performance, and frequent reviews are expected.

The nature of a formative assessment task need not differ if the task was to be a summative one (Scriven,1991). For instance, asking a student to write a use case description can be done as a formative or as a summative assessment task. The difference is that the formative assessment task serves the function of improving the students skill at writing use case descriptions whereas the summative assessment task is used to establish how well the student can write a use case description. The formative assessment tasks used in this study were taken from the summative assessments used in previous deliveries of the subject. The previous summative assessments, which contained many tasks, were broken down into individual tasks suitable for formative assessment.

For a student to be able to improve they need to be given the opportunity to repeat the formative assessment task (Bloom, *et al.*,1971), just as a developer will improve their analysis and design by iterations of doing and reviewing the task. Therefore, the design of the structured learning experience offered several opportunities for the student to repeat the same task.

The students would be experiencing such a rigorous structure of tasks for the first time. To address the change in the type of student as identified by Biggs, to enforce the structure of the learning strategy, and to let the student know that each task was considered important it was decided for this study to make the entire set of assessment tasks count toward a student's final grade. There was a risk that this decision would over-emphasise the grading and under-emphasise the learning as identified by (Black,2000). However, the prevailing attitude of students toward assessment was that unless the task counts directly to the final grade the student was not going to do the work. Also,

the formative assessment needs to prepare the student for summative evaluation (Scriven,1991).

Based on the literature the formative assessment strategy was a judicious teacher/responsive student feedback loop, feedback being a grade and comments; provided multiple opportunities to iterate the assessment task; provided opportunity to modify the instruction; emulated the work environment in which the knowledge and skills would be used; and was a shell into which effective learning strategies could be fitted (Black,2000); would "enhance learning by structuring the study effort" (Boud,1995); and was a repetitive, ongoing process, using multiple simple assessments to evaluate student learning (Boyd,2001).

3. The Subject

In the subject, Business Information System Development 1 (BISD), chosen for study, students learn a limited set of modelling tools and techniques that are used in the development of business information systems (Box and Ferguson,2001).

Formative and summative assessments were used as the means for students to acquire and demonstrate their knowledge and skill of object-oriented information system development modelling tools and techniques. The subject was conducted with two instructors delivering to 93 students. During the 12 weeks of teaching, there were 12, two-hour, weekly classes with all the students attending. Also, groups, of no more than 20 students attended one tutorial of one hour per week. The students also arranged individual and small group consultation with an instructor.

The specification of the subject was given in the subject outline. The subject outline is a document describing the aim(s), objectives, assessment, marking criteria, content, and teaching and topic schedule of a single subject. The previous subject outline had been identified as needing improvement in a number of areas to bring it in line with leading educational research findings and the extensive use of formative assessment.

4. The Subject Objectives

The objectives were formulated based on: previous subject outlines, teaching experiences with BISD and teaching experiences with follow on subjects, peer review, the desire to integrate generic attributes or transferable skills and course design literature (Biggs,1999, Boyd,2001, Crooks,1994, Boud,1995, Bloom and Krathowl,1956). These influenced the choice of the breadth of content, the desired depth of understanding of the content and, upon reflection, assured the quality of the objectives. The phrasing of the objectives was based on Bloom's taxonomy of cognitive abilities expressed in behavioural terms (Bloom and Krathowl,1956). The objectives were expressed as demonstrable outcomes achieved under certain conditions to a particular standard. This ensured, according to the literature, the assessment could be brought into line with the objectives.

The objectives, as they appeared in the subject outline, appear in Figure 2. The breadth of coverage stated in the objectives was narrowed from previous deliveries. It was necessary to do this to allow time for the students to achieve an adequate level of mastery in the most important areas of BISD. Reducing the content meant a limited set of modelling tools and techniques were taught. This limited set of models was used to perform the analysis and design of a system, without any loss of traceability from one model to another, and ended with a set of models that would be adequate to use as program specifications.

There was a shift from norm referenced to criteria referenced grading. The new objectives were grouped to indicate what a student would have to be able to do to achieve a particular grade. It was realised that the majority of the objectives had to be met for a student to pass. It was no longer workable to consider that 50% measured an adequate mastery of the essential materials. This meant that the previous method of distributing portions of 100% to various tasks, awarding marks for each task and then aggregating the marks to determine a grade needed to change. The change meant a student's grade was determined explicitly by his/her demonstrable achievement of all of the objectives for a particular grade. The use of the judicious teacher/responsive student feedback loop, the profile of pass/fail indicators a student would have, and the summative assessment was the means of determining a student's final grade.

5. The Formative Assessment Strategy

The assessment tasks were described generically in the subject outline. Specific content of each task was set week-by-week based on the progress of the students. The instructors gauged progress during class, at weekly meetings, by weekly marking of assessment tasks, from written student evaluation, and verbal and written unsolicited student feedback.

The formative assessment regime covered a topic over three weeks, weeks A, B & C in Figure 3. In class in week A, the topic was set. A student had from the end of class in week A until the beginning of class in week B to write a two-paragraph summary of the set topic: "lecture preparation". During class in week B, the topic was discussed and one or more exercises were completed: "lecture exercises". Between classes in weeks B and C the student attempted the exercise again: "tutorial preparation". During the tutorial in week C, a student made notes on his/her tutorial preparation; tutorial preparation plus these notes was the "tutorial participation". A student could then attempt the work again, "resubmission", after reflecting on his/her tutorial notes, pass/fail indicator and comments. The instructor provided grades and comments online, usually within 24 hours after each class in weeks B and C. This assessment regime meant, once the students were fully engaged with the subject, that they would be studying a number of topics a week and that they would encounter a topic at least five times over a number of weeks.

Aim:

For students to become adept with the knowledge and skills required to competently use a limited set of the tools and techniques within a given method of business information systems development.

Objectives/Learning Outcomes:

By the end of this subject, a student with at least a passing grade should be able to:

1. Recognise probable business requirements for a given scenario.
2. Accurately identify and specify manual and automated processes for a given scenario.
3. Write clear and unambiguous specifications at a business oriented high level and at business and system oriented expanded levels.
4. Identify members in a development team and their responsibilities while contributing to the processing and completion of the development tasks covered in this subject.
5. Recall and explain a software development process and the activities within the phases of the process covered in this subject.
6. Accurately represent a business information system as a set of models using common and well-established modelling tools and techniques.
7. Explain and repeat the process used to create a set of models that accurately represent a business information system.
8. Follow a self-evaluation process evidenced by a written product.

By the end of this subject, a student with at least a credit grade should be able to:

9. Do all the above objectives.
10. Independently improve on the set of models prescribed in class by responding appropriately to comments from others and his/her self-evaluation.

By the end of this subject, a student with at least a distinction grade should be able to:

11. Do all the above objectives.
12. Independently expand the set of models from those prescribed in class to represent a business information system for a larger scope of the given scenario.

By the end of this subject, a student with a high distinction grade should be able to:

13. Do all the above objectives.
14. Demonstrate professionalism in the completeness and accuracy of his/her work such that it would be pedantic to find fault with the work.

Figure 2: Aims and Objectives of BISD

A sample of four weeks from the teaching sessions, Figure 4, shows the topics covered, how doing the task for a following topic could help a student improve the proceeding work on a topic. The cascade of topics imitates the work environment. One artefact may be produced, another commenced, and this leads to discoveries about the developing system that causes the developer to revise the proceeding artefact. The formative assessment strategy emphasised the iterative and incremental software development approach.

The consequent, required, study effort was very structured with multiple simple assessments in a week, namely: one lecture preparation; one or more lecture exercises in class, tutorial preparation, and tutorial participation in class.

6. The Assessment Tasks

The judicious teacher/responsive student feedback loop required that the teacher set the task and criteria for evaluation of the task. Each assessment task was set by the teacher and a list of criteria for each task provided.

Assessment Tasks							
	Lecture Preparation	Lecture Exercise		Tutorial Preparation	Tutorial Participation		Resubmission
Week A (in class)	topic is set for next week						
Week B (in class)	topic is discussed, lecture preparation reviewed and submitted	exercise(s) on this week's topic are worked and submitted		the scope of the tutorial preparation is set			
Week B (after class)			lecture exercise(s) grades and feedback released online				
Week C (in class)				submitted (the tutorial preparation and participation are on the same paper)	preparation is discussed, a student makes notes on his/her tutorial preparation & then it is submitted		
Week C (after class)						tutorial preparation and participation grades and feedback released online	
Week C – end of week						students resubmit tutorial work	
							grades released prior to next class

Figure 3: Formative Assessment Regime for a Topic

Lecture preparation required a two-paragraph summary of a topic based on a few pages of reading. The preparation was similar to the “One-Sentence Summary” exercise-who does what to whom, when, where, how and why (WDWWWWHW) (Huba and Freed,2000).

Lecture exercises were usually a student’s first attempt at a model. The teachers discussed the model and what is needed to develop an accurate model. The students responded by developing all or part of the model. The teacher collected the work to judge it. Tutorial preparation

was a second attempt at the same model with a slightly larger problem scope. Tutorial participation was an opportunity to perform in class evaluations on the prepared model and make a third attempt when time permitted. A resubmission was the fourth attempt at the same modelling exercise.

A learning portfolio was the only summative assessment. The learning portfolio contained all attempts at a model, that is, all set tasks plus any more attempts a student wished to do. Each new attempt was made after evaluating

Assessment Tasks						
		Lecture Preparation	Lecture Exercise	Tutorial Preparation	Tutorial Participation	Resubmission
Week A (in class)	Topic set: high level use cases					
Week A to B (between classes)		Read & summarise: high level use cases				
Week B (in class)	Topic set: use case diagrams		Discuss & 1 st Attempt: high level use cases			
Week B to C (between classes)		Read & summarise: use case diagrams		2 nd Attempt: high level use cases		
Week C (in class)	Topic set: class diagram		Discuss & 1 st attempt: high level use cases		Evaluate & 3 rd attempt: high level use cases	
Week C to D (between classes)		Read & summarise: use case diagrams		2 nd attempt: use case diagrams		4 th attempt: high level use cases
Week D (in class)	Topic set: CRUD analysis		Discuss & 1 st attempt: class diagram		Evaluate & 3 rd attempt: use case diagram	
etc

Figure 4: Sample of four weeks of topics

the preceding attempt. The portfolio was a record of the student's change in knowledge and skills. The portfolio was a "purposeful collection of a learners work... [documenting] achievements, efforts, progress... [and] self-assessment (MacIsaac and Jackson,1994). The summative assessment task was completed by the students engaging in the advisory teacher/judicious student and/or self-determining student feedback loops.

The content, the topic to be summarised or the model to be attempted, were specified in the subject outline. The specification was modified as the weeks progressed in response to the learning and understanding shown in the submitted assessments. For instance, weeks 8 and 9 were a review to consolidate and clarify all preceding topics and clear up misunderstandings about the requirements of the system being designed, rather than the delivery of more content.

7. Method of Instruction

The unit was presented as a guided experiential learning environment. A role-played case (Cope and Horan,1996) of a genuine system was used. One of the instructors was the real user and client for the system. The students were supplied with a written case, several client briefings incorporating actual artefacts and photographs, and detailed responses to questions asked of the client. Students were able to question the client in class, via e-mail, and by out-of-class interview.

The students were encouraged to adopt appropriate study strategies to avoid severe workload problems from the outset (Crooks,1994). For instance, the lecture preparation required the reading and summarising of a few pages directly related to the topic rather than one or two chapters. The WDWWWHW technique was explained and emphasised during classes to help students keep to the two-paragraph limit.

Classes were conducted using a number of group activities interspersed with short disseminations of content and instructions. Questioning and interaction during classes was encouraged. Each class had a tangible outcome, a lecture exercise of one or more full or partial models, or evidence of tutorial participation by notes made on the tutorial preparation brought to class.

To facilitate student-teacher communication out of class a

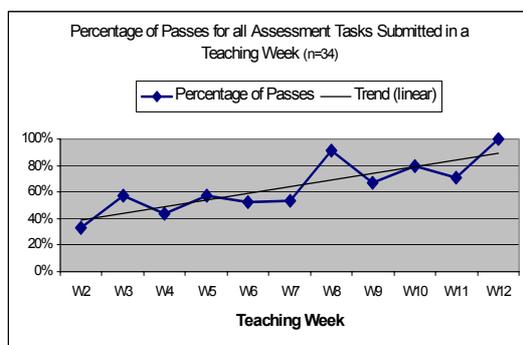


Figure 5: Percentage of Passes for Continuous Assessment Tasks on a Week by Week Basis

web-based course management tool, Blackboard 5, was used to 1) release grade sheets and feedback on submissions in a timely manner; 2) broadcast answers to students' e-mailed questions; 3) make announcements about administrative issues, and 4) provide materials such as sample source case documents, readings and class notes.

8. Results

Data collection for the study occurred by participant observation, i.e. by an instructor, conversations with the other instructor and students, judging assessments, a formative evaluation, and in-class polls.

The retention of students followed a typical pattern experienced in similar subjects at the same institution. At week one there were 93 students enrolled in BISD, by week six fewer than 50 were regularly attending, in week 12 only 34 students remained. In-class polls, disclosed by students in contact with others not attending classes, that the students who stopped attending and/or submitting had essentially given up and were resigned to failing. The following results relating to grades consider only the 34 students that completed, that is, only those that submitted a portfolio.

The percentage of passes the students received each week tended to increase over time, Figure 5. The correlation of passing formative assessments and passing the summative assessment is shown in Figure 6 and Tables 1 and 2. The distribution of grades for the summative assessment, Table 3, shows 38% of students failed and 62% received a pass or higher. Of the 93 students that commenced only 23% passed, of the 77 enrolled in week six 27% passed.

A formative subject evaluation was done in week 5, using the instrument included in the "Formative Evaluation of Teaching Strategy" in Hogan (1999). Table 4 shows the compilation of the responses into positives, areas of difficulty and suggested improvements. The total number of evaluations was 39. However, the five constructs in the instrument gave a student the opportunity to express his/her concerns more than once. All instances of a response by a student were included, hence the number of responses to "Explanation" under "Suggested Improve-

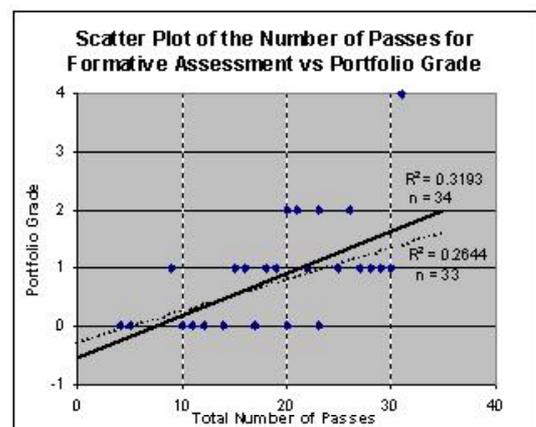


Figure 6: Number of Passes for Continuous Assessment vs Portfolio Grade

	Lecture Preparation	Lecture Exercise	Tutorial Preparation	Tutorial Participation	All Continuous
Lecture Ex	0.56				
Tut Prep	0.40	0.44			
Tut Part	0.38	0.43	0.72		
Portfolio	0.26	0.32	0.60	0.55	0.59

Table 1: Correlation Coefficients. Numerical variables are: mean number of passes for continuous assessment, and pass (1) or fail (0) for the portfolio (n = 34)

	Lecture Preparation	Lecture Exercise	Tutorial Preparation	Tutorial Participation	All Continuous
Lecture Ex	3.52				
Tut Prep	2.36	2.63			
Tut Part	2.23	2.56	5.06		
Portfolio	1.48	1.85	3.86	3.44	3.77

Table 2: Significance of Correlation Coefficients. (z scores) n = 34

Grade	No.	%
High Distinction	1	3%
Distinction	0	0%
Credit	5	15%
Pass	15	44%
Fail	13	38%
Total	34	100%

Table 3: Distribution of Grades for the Portfolio

ments" is greater than 39.

Another in-class poll, taken in the last class, collected data on the number of hours per week devoted to the subject. Early in semester, the instructors explained that about five hours out of class plus three hours in class would be needed to be applied to BISD to earn a passing grade. Only one student admitted spending eight or more hours per week, for every week of teaching. Approximately half admitted spending eight hours per week from week seven to the end of the teaching weeks. All of the students admitted spending more than eight hours per week on BISD in the last two teaching weeks.

During the 36 hours of individual or small group consultation, student comments were often volunteered or solicited. Typically, students brought work to the consultation. Students stated that their poor work was a result of not attending, not reading the class handouts, and/or not reading the online materials. Students also offered the comment that the structure of the assessment method gave them more opportunity to improve.

An instructor observed that the higher the grade for the portfolio the more likely the student had engaged in deep learning from early in the semester. Indicators of deep learning were that tasks were done correctly and met the requirements of the client. Surface learners tended to do the task without demonstrating an understanding of the correct way to do the task or of the client's requirements.

Formative Evaluation at Week 5 Results as Number of Responses (n=39)

Category	Exemplars of Responses	Positive	Area of Difficulty	Suggested Improvement
Environment	availability of consultations, access to the online materials, the client being readily available, the tutorial instructor and instruction, the administration setup	16		
Explanation	explanations of the assessment and topics, provide more diagrams, examples and answers to tasks, provide more system requirements, understanding what is expected, quantity of materials		27	52
Feedback	comments and response to each task	2		11
Group Work	collaboration with students during class, interaction with the client and instructors	18		
People	student whining, humour, jokes and banter in class, stupid questions		10	
Relevance	attending class is worthwhile, applying what is being taught/learnt	17		
Structure	repetitive structure of assessment, improving time management, having to do lecture preparation, practicing skills, strict policy of having to do tasks	27	35	
Technical	computer/printer failure		1	
Workload	the amount of preparation and reading required, sacrificing other subjects, not having enough time		14	

Table 4: Compiled Week 5 Formative Evaluation Results.

Surface learners used the notation for the models with little indication they had thought about the task. Some took a deep approach to learning the correct way to do the task, but a surface approach to understanding the purpose of the model, and so incorrectly represented the client's requirements. Many started out by taking a strategic approach, often querying their status, the number of passes and fails they had received, and how many more passes were needed. There were two or three asking these questions late in the subject.

9. Discussion

The results show that the goals: to engage students in deep learning; increase the learner's responsibility for his/her learning by participating in their own achievement of the objectives; and develop better study practices that are more

like a software development work environment have to some extent been achieved.

The drop in retention from 93 students to 77 and then 34, while concerning, is in accord with other subjects with mostly the same cohort, and with previous BISS cohorts. Since other, more traditionally delivered and assessed subjects, with mostly the same cohort, experienced similar attrition it cannot be concluded that the changes implemented for this subject are attributable to the attrition. The distribution of grades for the portfolio, Table 3, is much as expected based on observed and recorded student participation. Though more than half the students that submitted portfolios received a pass or higher it is disturbing that less than a quarter of the initial student enrolment persisted and achieved the objectives. For those that persisted, the results suggest some degree of achievement of the goals.

As students became familiar with the formative assessment strategy their results improved, Figure 3. Though the correlation coefficients, Tables 1 and 2, are low, all but two correlations, portfolio versus lecture preparation and exercise, are significant. The significance of the correlations is not surprising. It shows what most teachers already believe, that the more a student does to learn, the better the student's result. This is also indicated by the strength of the correlations shown by the R^2 values in Figure 6. These results are encouraging and support the statement that there was a degree of achievement of the goals.

The formative evaluation, Table 4, revealed a number of issues. The greatest number of responses was in the "Explanation" category. Though students found it difficult that there were no "right" answers, it was evident that more examples were required. This might be an indication that further improvement is required of the teaching by "provide concrete examples of ... abstract aims" (Black,2000).

The formative assessment strategy, Table 4 "Structure", was well received by the students although many were experiencing difficulty. Also, there was a positive response to the relevance of attending class and what was taught, "Relevance". This indicates that better study practices were encouraged.

The "Environment" and "Group Work", Table 4, received positive responses. The students were benefiting from interacting with the instructors and other students, and were engaged in learning.

There is a difficulty with explaining the data for workload when juxtaposed with the poll results of time spent on the subject. The one student that had applied eight or more hours throughout the session was astute and realised she required such an effort to learn. Others, perhaps not as astute, account for the contradiction.

There was not sufficient constant reinforcement of the connections among the formative assessment tasks, how it was preparing the students for summative assessment and the achievement of the objectives. The evidence being: the early strategic approach taken by many that persisted for a few; the failing summative assessment did not address all

of the objectives; the increased number of study hours in the final weeks.

The reasons for poor work as stated by the students show that some were taking on the responsibility for their learning and were responding to the judgment of their work appropriately and positively.

The observation that the deep learners were the 62% of passes for the portfolio compares favourably with 53% from a similar study (Cope and Horan,1996).

The learning approach taken by students is possibly a function of the feedback they received. It is suggested, that comments accompanied by grades do not lead to learning gains (Butler,1988). The threat the grades appear to have posed could account for the low number of higher grades and the attrition. Though the students received timely and detailed comments, the use of pass/fail indicators could have remained their focus of attention. Any gain that the instructor hoped to achieve by detailing the positives and suggestions for improvement for a task were, possibly, for the most part ignored. On one occasion a student commented, "There's not enough feedback", yet over six pages of comments on the previous weeks tasks had been provided. It may also be that the feedback was not effective.

(Crooks,1994) states the requirements for effective feedback as:

1. "... truly informative, clearly identifying ... strength[s] and weakness[es] of the individual student."
2. "ways of improving ... should be ... explicit as possible."
3. "delivered ... to enhance rather than damage student motivation."
4. "timely."

The instructors were aware of these requirements and believed they had been met. Upon reflection, perhaps the missing element was that the comments were not to an individual, all comments on a task were compiled into one document, broadcast to the students and each student was expected to correct the comments to his/her work and respond appropriately. The difficulty in meeting this requirement would cause a significant increase in the instructor's workload, as mentioned earlier this is a likely weakness of the judicious teacher/responsive student feedback loop. To assuage the load of individualised comments the instructors grouped the comments. For many students the lack of individualised comments may have broken the feedback loop, shown by many students choosing not to participate. An alternative may be to improve the students' responsiveness to comments and their evaluation of their work. From the evidence of this study, for many students comments accompanied by grades do not lead to learning gains.

The debate about the difference between formative and summative assessment has received attention in recent years (Kanjee,2000). The formative assessment tasks, which were designed to encourage learning and student development, would be considered as formative (Kanjee,2000, Black and William,1998, Cowie and

Bell,1999). However, most students commenced the semester with the position that the assessment regime was "a multiple set of summative assessments... with a considerable direct influence on the final outcome" (Bligh,2000). For some, this position persisted for the semester. Though the assessment was continuous, perhaps it was not formative, that is, providing a non-judgmental, non-threatening atmosphere (Rolfe and McPherson,1995)

5. Conclusion

This study has been beneficial. The use of a rigorous and structured formative assessment strategy yielded positive results. The successful students did engage in deep learning, increased their responsibility for their learning, and improved their study practices.

A purpose of the evaluation of BISD was to generate ideas for further improvement. Some ideas are: 1) in the future, each student may decide when they would like feedback only or feedback and a grade for a task and so make the atmosphere less threatening and judgemental; 2) the portfolio will receive a number of formative evaluations during the semester thereby formalising the advisory teacher/judicious student feedback loop; 3) the work the students do to complete a task will be reviewed as well as the final outcome of a task; 4) the connections among the objectives, each task and the portfolio will be emphasised; 5) increase the provision of self assessment techniques to make better use of peer and instructor feedback.

6. References

- BIGGS, J. (1999) *Teaching for quality learning at university* Buckingham, England, Society for Research into Higher Education and Open University.
- BLACK, P. & WILLIAM, D. (1998) Assessment in classroom learning, *Assessment in Education*, 5, 7-71.
- BLACK, P. (2000) Research and the development of educational assessment, *Oxford Review of Education*, 26, 407-419.
- BLIGH, D. A. (2000) *What's the use of lectures?* San Francisco, CA, USA, Jossey-Bass Publishers.
- BLOOM, B. S. *et al.* (1956) *Taxonomy of Educational Objectives Handbook I: Cognitive Domain* Longmans, Green and Co.
- BLOOM, B. S., HASTINGS, J. T. & MADAUS, G. F. (1971) *Handbook on Formative and Summative Evaluation of Student Learning*, 1-923.
- BOUD, D. (1995) *Implementing student self-assessment* ACT, Australia, Higher Education Research and Development Society of Australasia Inc.,
- BOX, I. & FERGUSON, J. (2001) *Improving Software Development: The Prescriptive Simplified Method*, *Australian Conference on Information Systems*, Coffs Harbour, NSW, Australia, December 5 - 7 , 2001.
- BOX, I. (2002) *Applying Educational Research to Improve Teaching and Learning in Information Systems*, Paper presented at the *Informing Science and IT Education 2002*, Cork, Ireland, 19 - 21 June 2002.
- BOYD, B. L. (2001) *Formative classroom assessment: learner focused*, *The Agricultural Education Magazine*, 73, pp18.
- BUTLER, R. (1988) *Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance*, *British Journal of Educational Psychology*, 58, 1-14.
- CHASE, C. I. (1999) *Contemporary assessment for educators* New York, Addison-Wesley Publishers, Inc..
- COPE, C. & HORAN, P. (1996) *The role played case: an experiential approach to teaching introductory information systems development*, *Journal of IS Education online*, 8, 33-39.
- COWIE, B. & BELL, B. (1999) *A model of formative assessment in science education*, *Assessment in Education*, 6, 101-116.
- CROOKS, T. (1994) *Assessing student performance* Campbelltown, Australia, Higher Education Research and Development Society of Australasia Inc..
- HOGAN, C. (1999) *Facilitating learning: practical strategies for college and university* Melbourne, Australia, Eruditions Publishing.
- HUBA, M. E. & FREED, J. E. (2000) *Learner centered assessment on college campuses: Shifting the focus from teaching to learning*. Boston, Allyn and Bacon.
- KANJEE, A. (2000) *Investigating formative assessment: Teaching, learning and assessment in the classroom*, *Assessment in Education*, 7, 160-163.
- MACISAAC, D. & JACKSON, L. B. (1994) *Assessment processes and outcomes: portfolio construction*, *New Directions for Adult and Continuing Education*, 63-72.
- ROLFE, I. & MCPHERSON, J. (1995) *Formative assessment: how am I doing*, *Lancet (North American Edition)*, 345, 837-9.
- SCRIVEN, M. (1973) *The Methodology of Evaluation*, in: Worthen, B. R. & Sanders, J. R. (Eds.) *Educational Evaluation: Theory and Practice*, pp. 60-104 Belmont, CA, Wadsworth Publishing Company.
- SCRIVEN, M. (1991) *Beyond formative and summative evaluation*, in: McLaughlin, M. W. & Phillips, D. C. (Eds.) *Evaluation and Education: At Quarter Century*, pp. 18-64.
- TORRANCE, H. P., JOHN (2000) *Investigating formative assessment: Teaching, learning and assessment in the classroom*