Assessment for learning and accountability: an integrated approach

Shukri Sanber

School of Education, Australian Catholic University, Shukri.Sanber@acu.edu.au

The literature indicates that a key feature of assessment in practice is over-emphasis on grading and under-emphasis on learning (Black & William, 1998a, p. 18). This feature might be attributed to many factors, such as lack of understanding of the characteristics of formative assessment and the need to meet grading and reporting requirements. Achieving a balance between assessment tasks that support learning and those designed for the grading function is not easy. The characteristics of assessment items that help to achieve these two functions are not identical. This paper reports on the outcomes of a small-scale study that was conducted as part of a core unit in a teacher education program at an Australian university. Two lines of assessment were built into this unit. One was programmed to provide the students with ongoing and detailed feedback; the students were encouraged to use this feedback to improve their performance. The other was made up of tasks designed for grading and reporting purposes. Students' perceptions and comments were collected through a questionnaire and focus group interviews. The performance of the participating group of students of the various tasks was tracked. Data analyses indicate that the assessment experiences of this group of students improved their final performance, increased their awareness of their own learning and helped transform their learning into a deeper understanding of the focus of the unit and its role in their future careers as teachers.

Keywords: accountability; assessment for learning; formative assessment; student feedback, summative assessment.

Introduction

Workload and assessment are two areas that students in higher education complain about. Results of student feedback gathered through unit experience questionnaires confirm that students do not always rate assessment and workload as favourably as other unit dimensions such as organisation, resources and content (Ramsden, 2003; McInnis, James & Hartley, 2000; Sanber & Hewitt, 1995). There is no doubt that assessment plays a pivotal role in the process of learning: it is the focus of attention at all levels of education. Hence, educationalists call for better integration of assessment into learning. For example, Ramsden (2003) argued for a "view of assessment as being (a) a means of helping students to learn, (b) a way of reporting on student progress and (c) a way of making decisions about teaching" (p. 205). Assessment serves both the student and the teacher. The formative role of assessment is justified and propelled by the movement towards assessment for learning. A great deal of emphasis is given to formative assessment in educational literature, because it stimulates student learning (Black & William, 1998a; Nicole & Macfarlane-Dick, 2006; William, 2006). Teachers at colleges and universities as well as at schools are encouraged to use ongoing assessment and to provide students with detailed feedback (Gibbs & Simpson, 2004).

Assessment policies at many universities both in Australia and overseas are increasingly prescriptive in outlining assessment procedures. Failure to follow these prescribed procedures might be used as grounds for student appeals against awarded grades. Lecturers are expected to communicate to their students the required assessment tasks at the beginning of each semester (The Australian Catholic University, 2009, p. 57). These assessment tasks are usually used for the purpose of grading; thus, they have a principally summative function.

Achieving a balance between formative assessment and summative assessment is not easy. Since formative assessment is designed to support student learning, which is dynamic and changing, assessment procedures and strategies designed for that purpose are expected to be adaptable and not strictly predetermined. On the other hand, assessment of learning that serves accountability requirements needs to be structured and predetermined. The lower

ratings given to assessment in unit experience questionnaires may indicate that the balance between formative assessment and summative assessment is not always being maintained. It is likely that more emphasis is being given to the latter purpose of assessment.

The terms 'formative' and 'summative' were originally used to differentiate between two purposes of evaluation (Scriven, 1967). Formative evaluation takes place before and during the implementation of the program, "with the intent to improve it" (Scriven, 1991, p. 169). Summative evaluation is conducted with the purpose of "reporting on it" (p. 340). As assessment was differentiated from evaluation in the 1980s under the influence of the 'authentic assessment' movement, the terms formative and summative assessment started to appear in the literature (see, for example, Sadler, 1989). Formative assessment is carried out as learning takes place to guide the learner and the teacher. Summative assessment takes place at the end of the program of learning and is used mainly for grading and reporting.

Assessment is an information-dependent process. The information can be numerical or descriptive, and it can be gathered formally or informally. Assessment tasks used for formative and summative purposes differ at least on the basis of two dimensions (see Figure 1): namely, the nature of the domain of assessment, which may range between specific and broad, and the degree of emphasis on differentiating performance levels, which may range between low and high.
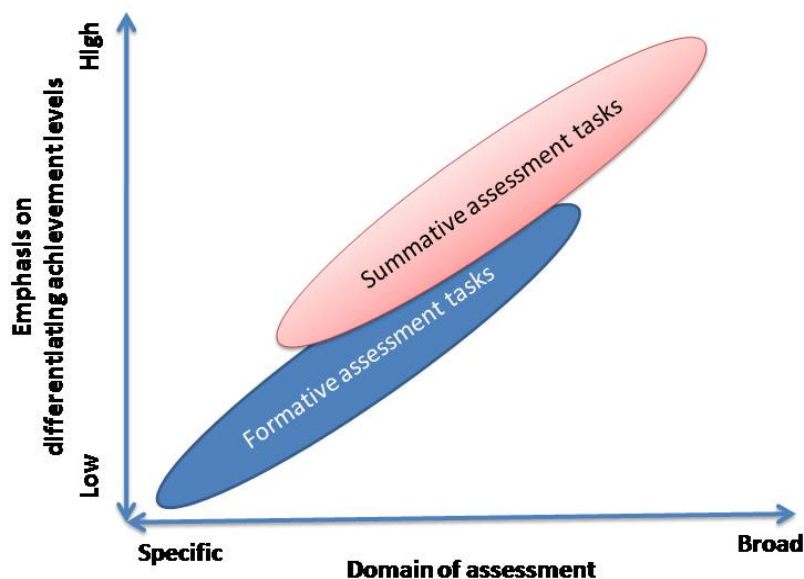

Figure 1: Nature of tasks used in formative and summative assessment

The characteristics of assessment tasks used for formative purposes cannot be identical to those used for summative purposes. Formative assessment tasks are designed to facilitate student learning (William, 2006): they need to target specific domains, and their items are expected to exhibit a wide range of difficulty indices, otherwise they will not be effective in generating corrective feedback to help student learning (Hattie & Timperley, 2007). The items or components that make up assessment tasks designed for summative purposes need to represent a wider range of the domain of assessment in order to be valid. They do not target specific information but rather encompass the breadth of the domain of assessment. Such tasks are usually designed to maximise the spread of students' scores to facilitate the process of grading (Miller, Linn & Gronlund, 2009).

Contemporary researchers in the area of assessment promote formative assessment and the role of timely feedback. Yet the literature on assessment in higher education seems to ignore the fact that tasks used for formative purposes are not identical in their characteristics to those used for summative purposes. Furthermore, there seems to be no

evidence to show that it would be feasible to design assessment programs that combine sets of tasks dedicated specifically to formative purposes and parallel sets of tasks dedicated to summative purposes in the same units.

This paper reports on the outcomes of a small-scale project in which both formative and summative assessment tasks were incorporated into the same unit. The project was designed to meet the purposes of supporting student learning by generating specific and individualised feedback as well as attending to the requirements of grading and reporting.

Integrated assessment project

The project was implemented over one semester in 2008. It was carried out within the context of a core unit in a secondary teacher education program at an Australian university. This core unit is normally offered to fourth-year students.

The major assessment component assigned in the unit was a 'test construction project' (TCP). The students were required to choose and analyse a unit of work based on the school curriculum in their area of specialisation, and then construct a test to assess the achievements of their future students. The process of test construction was divided into three stages, with each stage comprising an assessment task (see Figure 2). Students were required to complete and submit these tasks on specified dates during the semester. The tasks were aligned with the key concepts being discussed in the unit. In addition to these three tasks, students were asked to submit two 'in-depth tasks' in which they were to identify and discuss issues related to assessment in their area of specialisation. The in-depth tasks were designed to inform the planning and implementation of the TCP.

The lecturer made sure the tasks were returned to the students within one week of submission, along with detailed and individualised feedback. Marks were not awarded to these ongoing tasks. The students were informed at the beginning of the semester that these tasks were designed to support their learning. They were encouraged to reflect on the feedback and to incorporate relevant components into the final submission of their project. They were informed that the completion of the ongoing tasks was not optional, and that a copy of the feedback would be stored in their unit file.
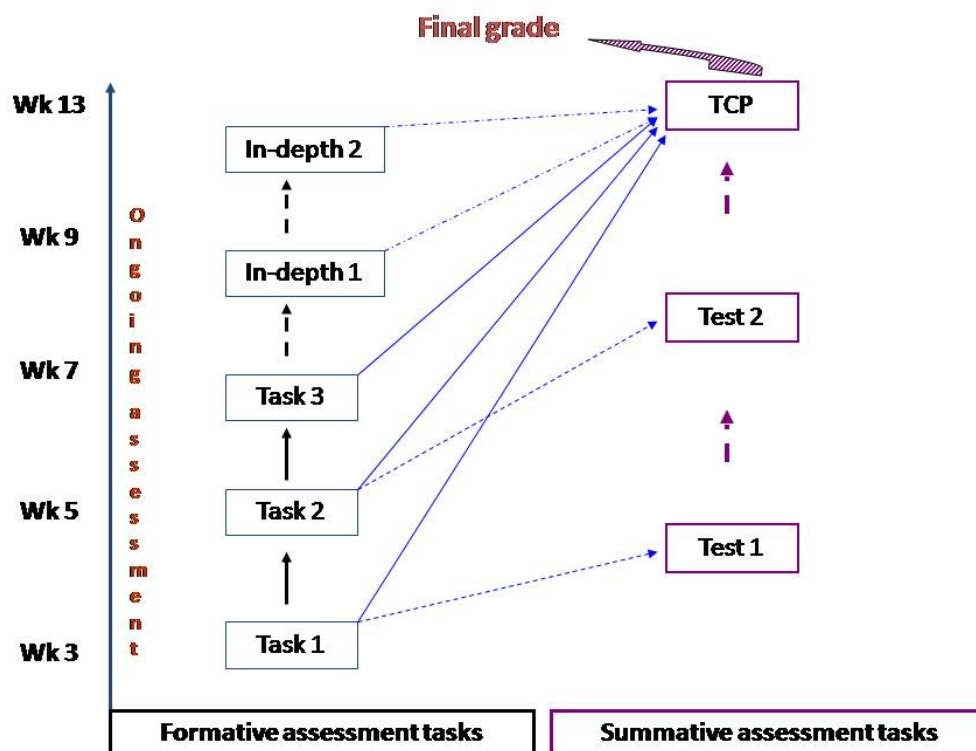
Figure 2: Assessment schedule during the semester

To serve the summative purpose of assessment, two tests were administered, in Weeks 4 and 8 of the semester (see Figure 2). Each test comprised 50 multiple-choice questions. The tests covered the assigned material for Weeks 1 to 3 and Weeks 4 to 7, respectively. The questions were constructed to represent the full breadth of the assigned material. Their difficulty indices ranged between 30 and 65 to ensure a reasonable spread in students' performance levels and therefore facilitate grading. The tests were independently reviewed by an experienced teacher for language, clarity of expression and content validity.

At the end of the semester the students submitted their TCPs. These were marked against a set of criteria that had been made explicit to the students at the beginning of the semester. The final grades were based on the students' accumulated performance on the TCP and the two summative tests.

This paper reports on the results of the evaluation of this innovative assessment project, which integrated two lines of assessment tasks to serve two different purposes of assessment. The assessment items or components that formed each line differed in their breadth and in their distribution of difficulty indices to ensure that they met the summative and formative purposes of assessment, respectively. The main purpose of this paper is to report on the feasibility of implementing these two lines of assessment. The paper specifically addresses the following questions.

1 What were the features of the integrated assessment project as perceived by the students?

2 What were the factors that facilitated student learning during the implementation of the project?

3 What was the impact of the integrated assessment project on student learning, as measured by their performance on summative tests?

Methodology

The study was designed using a mixed-methods approach. The nature of the project, and the fact that it was implemented as part of a core university unit, necessitated the gathering of quantitative data through well defined assessment procedures. At the same time, the value of the outcomes of such initiatives depends on participants'

experiences. To this end, it was necessary to gather directly the perceptions of the participating students and to identify the meanings they derived from their own experiences. Therefore, qualitative data-gathering approaches were employed. This mixed-methods approach helped in acquiring evidence on the value of the integrated initiative from multiple perspectives, which are triangulated in this paper.

Participants

The total number of students participating in the study was 100. They majored in mathematics, technology, humanities (history or English), religious education, theology or visual and creative arts. All were completing their fourth year of study, and all had completed four weeks of practical teaching experience prior to enrolling in this unit. It is possible that some of them would have had some practical experience of educational assessment prior to enrolling in this unit, but none would have had any formal training in assessment.

Data sources

Three data sources were used in this project. The first was students' marks on the two summative tests and on the TCP. The two tests were machine-scored, while the TCP was first marked by a teaching assistant and then reviewed by the researcher. The marking was based on a predetermined set of criteria that had been made available to the students prior to final submission of the TCP. The marker and the reviewer established a common interpretation of the criteria and identified work samples that would represent each performance level prior to the marking.

The second data source was a 12-item questionnaire constructed to collect students' experiences of assessment in the unit. The questions were constructed to represent the following constructs: domain of assessment; relationship between assessment and learning; and quality of feedback received by the students. The internal reliability (Cronbach's alpha) of the questionnaire was 0.899. The questionnaire was administered at the end of the last session of the semester. The participation rate was 96 per cent.

The third data source was the discussions of two focus groups. The focus group discussions were held at the end of the university's formal examination period but before the grades were announced. Three students representing each of the six major specialisations of the participants (a total of 18 students) were invited to reflect on their experiences of assessment in this unit. Sixteen students initially agreed to participate in the discussions, of which 11 actually did. They were assigned to two groups: one for those who had majored in mathematics or technology, and one for those who had majored in humanities, religious education, theology or visual and creative arts. The discussions were facilitated by the researcher while a research assistant observed the discussions and took notes. The sessions were also audiotaped with the participants' permission. The following questions were presented for discussion.

1 What was the purpose of assessment in the unit from your point of view? Was this purpose achieved? How?

2 How did your experience in this unit influence your learning?

3 How did the various assessment tasks in the unit shape your learning experience?

4 Has the assessment experience in this unit influenced your understanding of the curriculum of your major area of specialisation? How?

5 Describe your views on the assessment workload in this unit.

Data analysis

Both the quantitative and qualitative data was stored in Microsoft Excel worksheets. The qualitative data was coded, grouped and categorised, and the resultant categories were analysed thematically (Creswell, 2008) to produce the major themes described below that summarise the participants' input.

Results

Features of the integrated assessment project

The main features of the project as derived from the students' input were as follows.

- **The ongoing assessment and feedback established grounds for improvement**
  The participants observed that they were assessed as they were learning the topics. The fact that feedback was provided within a week helped develop their positive perceptions. They had opportunities to use the feedback in later submissions. The focus group discussions showed unanimous agreement that the ongoing assessment established grounds for improvement. The following two quotes provide examples of students' positive perceptions.

  > I thought it was a very formative assessment, all of them, they were done at the same time we were doing that topic so it helped us really understand what we were doing – or even if we didn't when we handed it in and the feedback came back we then we could use that to further our knowledge (Student 2).

  > Some lecturers would mark it, not provide a marking criteria, no grounds from where you can improve or anything like that, and then you get the same mark for a similar assignment – it's like, well, obviously you didn't tell me where to improve. But with this one, with the TCP and how it was segmented, we were able to improve and apply the things that you actually suggested (Student 6).

- **The responsive feedback supported student learning**
  Support was made available through modelling, scaffolding and classroom activities. Feedback was designed to be responsive to students' needs, and was based on criteria that had been explained prior to submission of the tasks. It recognised good performance and identified areas of improvement. The students recognised that the ongoing assessment tasks were designed to support their learning. For example, Student 3 stated:

  > I thought that the way you constructed the TCP was really good – that it was formative, so we were learning as we went. I think that's really important, and we had a chance to have that scaffolded and then modelled for us, and then we went about doing this assessment.

Factors that facilitated student learning

The main factors that facilitated student learning during the implementation of the project, as derived from the focus group discussions, were as follows.

- **Relevance of the assessment tasks to student careers**
  The TCP and its assessed stages were designed to provide students with relevant contexts in which they would apply the knowledge and skills gained in the unit. Thus, the students selected and analysed concepts with which they were familiar and that they would most likely use as teachers.

- **Targeting transferable skills**
  The assessment tasks in this unit, including the summative tests, targeted higher-order thinking skills that can be transferred to other learning settings. The tasks required students to match, classify, integrate and apply. The focus group participants realised the value of these cognitive skills to other units in their academic programs. The ongoing assessment tasks provided them with opportunities to apply these higher thinking skills in concrete situations. These opportunities reinforced the value of the learnt skills in similar situations.

- **Use of individualised feedback**
  The feedback that followed the ongoing assessment tasks was individualised. It was presented in ways that helped the students to realise the relevance of the assessment tasks to them. This feature was one of the essential characteristics of the project. The feedback was based on common criteria, but provided advice and support in response to the input of the individual student.

The following quotes provide examples of students' positive perceptions.

Later on in the future I'm sure I'll use at lot more higher-order thinking questions in my tests, like a lot more 'create' questions and 'evaluate', whereas before it would've been lower, like 'identify' and … 'find x' and stuff like that (Student 1).

It just helped me in general [to] analyse the questions, like not necessarily just for tests but when we got surveys (I do surveys for money) and you read the questions and the way they phrase them and it is ingrained and you're analysing for any kind of leading parts to the question (Student 2).

The feedback was very helpful. This is something that really stuck out to me – getting the little letter after we had done the test with our marks on it, like that, I've never had that from a lecturer and I was like, oh wow, you valued and you took the time to sort of say thank you for doing the test, this is what you got, and you sort of broke down where we were, where we sat, how we did, and I thought that was really good (Student 3).

Impact of the project experience on student learning

Analysis of the data indicates that participating students achieved most of the outcomes of the unit. Firstly, their final grades (which were based on the two tests and the TCP) indicate that the majority achieved Credits (see Table 1). Based on past experience of this unit, and in accordance with university policy, the majority of the students should have been awarded a Pass. There is no doubt that the ongoing feedback and opportunities to make use of this feedback provided to the students improved their performances.

Table 1: Distribution of the grades of participating students

Grade	Percentage
Distinction	14
Credit	73
Pass	12
Fail	1

Secondly, the results derived from the two summative assessment tests (see Figure 3) indicate significant improvement in students' achievement between the two tests (t-test = 3.033; degree of freedom = 99; p = 0.003). These tests were designed for summative purposes. They were parallel in item characteristics and item type.
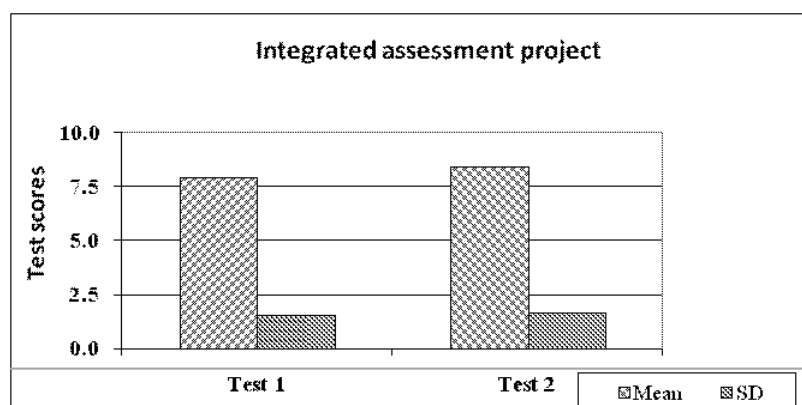

Figure 3: Change in students' scores on the two summative tests

Thirdly, students' input during the focus discussions, which were collected before their results were announced, indicates that they were confident that they had achieved the outcomes of the unit. This confidence in their performance can be explained by the close relationship between the learning and assessment processes in the unit. For example, one participant stated:

> You're actually learning about what you're being assessed on, and I think that, for me, was the real crunch – that not only is it practical but the stages we're going to be in was what was being assessed, and I think that's what I took from it (Student 5).

As well as helping them to achieve the unit outcomes, the participants' input indicated that their experiences in the unit helped them to:

- reflect on past learning

- transfer their knowledge and skills to other units

- improve their self-confidence.

Some of their comments were as follows.

> In previous units last year we were given the task of creating assessments, but we just did it so it looked like the real thing, but when we had to do this one I felt like it *was* a real thing, I felt like I could actually give it to a school and say, "Look, I followed the syllabus, this is what the curriculum is telling me to do" (Student 5).

> Just on how useful the learning has been, it was funny because we had an interview with the Department of Education the other day, and it was fantastic because one of the questions they ask you [is] "Well, assessment is an important tool, and how would you …" and I'm like, "I'm prepared for this," and I'm cracking out words like 'formative' and 'summative' and they're all like, "Oh yes," [so] I was happy about that (Student 6).

> Because I have an ad hoc approach to my learning, this subject forced me to apply myself a lot more, because I knew that I couldn't just 'go with the flow' – I had to really sit down and do this sort of work. Like with history or English it's different – it's completely different (Student 3).

The comments of the participants in the focus groups overall indicated that they had reflected on their past learning and experiences. They seemed to be aware of the value of the knowledge and skills gained in this assessment to other units and situations. The collected evidence indicates that students' assessment experiences in the unit helped them to examine critically their past learning and study habits and identify grounds for improvement. Finally, the

students' comments indicated improved levels of self-confidence. The students themselves observed that their experiences in the unit had helped them to be more realistic in their expectations and about their own capabilities.

Participants' perceptions of the assessment experience

The students completed several tasks during the semester. The majority of these tasks were not extensive, but they targeted higher-order thinking skills. The participants indicated that their workload was reasonable. The authentic nature of the tasks, as well as their relevance, their schedule and the support the students received, helped in forming their positive views. The following two quotes represent their perceptions.

> I think, for me, the workload wasn't too much. I know a lot of people have – well, all of us complain about uni all the time, but when I really think about what I've learned out of this unit, it's good how it was just all different. Like, you had tests and then you had a scaffold assignment, then you had the online then you had the comparison – it was all different, and in the real world you have to do that on a daily basis anyway (Student 4).

> I think there was a lot of work, but I think it was good that everything was spread out over two or three weeks. It didn't actually seem like that much. If you put all of the assignments together it definitely made up for an exam, because you did assess all of the topics, which was good, but it was very good that it was spread out. And how we did parts of the TCP in advance, so that towards the end in Week 9 and 10 it didn't really seem like that much work, like it did for the other units. So I thought that was good (Student 1).

Participants' ratings of their experiences during the project were analysed (see Table 2). The 'disagree' ratings shown here represent collated 'strongly disagree' and 'disagree' responses, while the 'agree' ratings similarly represent collated 'agree' and 'strongly agree' responses.

Student ratings are consistent with the results obtained from other sources. These ratings indicate that:

- the experience was relevant to students' future careers
- the feedback was responsive to students' needs and helped them to reflect on their submitted work and modify it accordingly
- the assessment tasks targeted a variety of outcomes that represented diverse content and cognitive skills.

The lowest rating was given to receiving informal feedback, although more than 56 per cent agreed with this statement. The students had the option of seeking conferences with the lecturer to discuss their progress. Many students made use of this option but not all of them did.

Table 2: Participants' evaluation of their experiences of the integrated assessment project

Experience	Disagree (%)	Undecided (%)	Agree (%)	Count (n)
In this unit I had the opportunity to modify my work	3.2	9.5	87.4	95
The assessment in this unit targeted a wide range of outcomes	3.1	8.3	88.5	96
The assessment in this unit targeted multiple forms of thinking	2.1	13.7	84.2	95
The assessment in this unit was integrated with the program of work	3.1	5.2	91.7	96
I used the feedback I received to go back over what I had done in my work	1.1	7.4	91.6	95
The assessment in this unit made me reflect on my work	4.2	5.2	90.6	96
I have used the skills or insight I gained in this unit with other units in my program of study	3.1	18.8	78.1	96
I feel that my experience in this unit has contributed towards my future career	1.0	7.3	91.7	96
The feedback I received prompted me to go back over material covered in this unit	7.3	18.8	74.0	96
I received informal feedback about my work in this unit	24.0	19.8	56.3	96
The number of assessment tasks and tests in this unit was optimal	25.3	12.6	62.1	95
I found the schedule of assessment (tasks and tests) in this unit helped my learning	6.3	14.6	79.2	96

Discussion and conclusion

This study was conducted in the first place to improve student learning in a core unit that includes complex concepts and specialised skills. To serve student learning effectively and to meet grading requirements, the researcher introduced an innovative assessment project with two parallel lines of assessment tasks administered progressively. The study explored the feasibility of dedicating separate assessment tasks to serve summative and formative purposes within the same unit. The first assessment line was designed to facilitate and support student learning through immediate, detailed and individualised feedback. The importance of immediate feedback is well established in the literature (see, for example, Gibbs & Simpson, 2004; Nicole & Macfarlane-Dick, 2006). The formative assessment line was truly an assessment for learning scheme, where students learnt through engagement and feedback. The assessment tasks within this line were designed to be specific and criterion-based, and to provide information to both the students and the lecturer. The assessment items had a wide range of difficulty indices so as to assess maximum performance. Students' responses were marked against public assessment criteria but were not graded. The most noticeable trend associated with this line of assessment was continuous improvement in students' performance.

The second assessment line was designed to help in grading decisions. This line was made up of two summative tests and a major assignment. The tests were administered at the end of the fourth and eighth weeks of the semester, respectively. They were designed to represent the breadth of the targeted content. The major assignment was designed to be a culmination of students' efforts throughout the semester. It was marked against a set of criteria that had been discussed with the students at early stages of the semester.

The participating students indicated that they had benefited from the learning and assessment activities implemented in this unit. Their final grades, which were based on the summative tasks, were higher than expected. The factors that facilitated learning in this study are consistent with those identified by Gibbs and Simpson (2004). The alignment between the targeted content and the ongoing assessment tasks, in combination with the immediate detailed feedback, helped students to reflect on their past and current experiences.

One salient feature of the successful implementation of the project was the disengagement of formative tasks from marks. Past experience has shown that marks distract students from reading feedback thoroughly. The summative nature of marks is not consistent with the nature of formative assessment. It should be mentioned here that during the initial stages of the project the students sought to 'work out' their marks, but as they progressed in the unit they began to focus on the feedback and on how to improve their performance on the next task.

Formative assessment tasks should not be used directly to grade students. They should be used to help improve students' learning. As the needs of students change with their progress in a unit, learning and assessment activities need to be redesigned accordingly. The number, format and focus of ongoing formative assessment tasks need to be flexible. Assessment policies should recognise the distinction between tasks designed for summative purposes and those designed for formative purposes. Lecturers should be encouraged to use truly ongoing formative assessment tasks that target specific domains in accordance with their teaching plans. These tasks should be treated as monitored learning activities. Grading and reporting should be based on the summative assessment tasks aligned with unit outcomes, but should represent broader course content and skills.

References

Black, P. & William, D. (1998a). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-73.

Black, P. & William, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Creswell, J. (2008). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (3rd ed.). Upper Saddle River, New Jersey: Pearson Education.

Gibbs, G. & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3-31.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Leahy, S., Lyon C., Thompson, M., & William, D. (2005). Classroom assessment: Minute by minute, day by day. *Educational Leadership, 63*(3), 19-24.

McGregor, K., Merchant, A., & Butler, M. (2008). Making feedback more immediate. *CAL-laborate International, 17*, 8-11.

McInnis, C., James, R., & Hartley, R. (2000). *Trends in the first year experience in Australian universities*. Canberra: Department of Education, Training and Youth Affairs.

Miller, M., Linn, R., & Gronlund, N. (2009). *Measurement and assessment in teaching* (10th ed.). Upper Saddle River, New Jersey: Pearson Education.

Nicole, D. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*, 199-218.

Ramsden, P. (2003) *Learning to teach in higher education* (2nd ed.). London: RoutledgeFalmer.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Sanber, S. & Hewitt, D. (1995). Student feedback using different kinds of questionnaires: How does it differ? In A. Brew, M. Jackson, J. Lublin, & M. Posser (Eds.), *Judging the quality of teaching: Approaches to university-wide student feedback* (66-81), Sydney: Centre of Teaching and Learning, University of Sydney.

Scriven, M. S. (1967). The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (American Educational Research Association Monograph Series on Curriculum Evaluation). Chicago: Rand McNally.

Scriven, S. M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park: Sage Publications.

Stufflebeam, D. L. (1983). The CIPP model for program evaluation. In G. Madaus, M. Scriven, & D. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 117-141). Boston: Luwer-Nijhoff Publishing.

The Australian Catholic University. (2009). *Handbook 2010*. Virginia, Queensland: Author.

William, D. & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Journal, 22*(5), 537-549.

William, D. (2006). Formative assessment: Getting the focus right. *Educational Assessment, 11*, 283-289.