Improving University Teaching and Learning

MIXED MODE ASSESSMENT: A PRELIMINARY EVALUATION

Douglas L Morgan

Yunggorendi

Abstract

This paper examines the usefulness of a combination of self, peer and tutor (mixed mode) assessment of group presentations for promoting deep learning and enhancing the learning experience of students undertaking the first year topic Aust 1004 – An Introduction to Aboriginal Studies, part of the Indigenous minor at Flinders University. It discusses the relative advantages and disadvantages of the three assessment processes and argues that a combination of all three can create an environment to promote deep learning. It evaluates the experiences of 25 first year students using this process and finds that deep learning is enhanced by mixed mode assessment.

Background

The aim of Aust 1004 – An Introduction to Aboriginal Studies is to give students tools that they can use for life. It is a first year topic which is a core subject in some courses (for example, Cultural Tourism) and an elective for others (such as the Bachelor of Arts - Australian Studies and Archaeology). As a result of undertaking this topic the student should be able to engage in debate about the diversity of Indigenous culture and issues, deconstruct and recognise the effects of racist and social theory on Indigenous Australians and be able to act as an agent for change. Students are expected to engage in the topic (demonstrated by essays and participation in tutorial discussion of current issues and their relationship to topic material) and be able to apply what they learn to novel situations. This requires not only that they gain information that can be retained and reproduced for evaluation (surface learning) but more importantly, that they can be used to integrate and reinterpret reality in a different way (deep learning) (Cartwright, 1997).

Deep and Surface Learning

Surface learning is encouraged by heavy workloads, prescriptive course materials, lack of choice within the topic, lack of choice in study methods, high topic contact hours and an assessment system that provokes anxiety. The process of learning is often perceived by the learners to be external to them and in extreme cases as something that is imposed by the teacher. In contrast, deep learning encourages the learner to understand the real world (Gibbs cited in Cartwright, 1997). It is fostered by the systematic involvement of the student in choice of topic content, allowing them to make connections with past learning, to plan their learning, reflect on the process and make abstract connections (Gibbs cited in Cartwright, 1997).

In Aust 1004 there are a number of barriers to promoting deep learning. Many of the students undertaking the topic do so because it is a requirement of their course Major. Some students are less committed, those who expected the course to focus on cultural differences have found racialisation theory confronting and some overseas students failed to understand the relevance of the course material to their own context. Another factor that may impact on the attainment of deep learning is the process of assessment itself. Gibbs (cited in Cartwright, 1997) found that forced compliance and/or a threatening anxiety provoking assessment system promotes surface learning. Many students find oral presentations in front of an audience the most threatening part of the topic, but since this is a requirement for passing the topic it can also impede deep learning (Gibbs cited in Cartwright, 1997).

While the very act of grading student performance may be counterproductive in producing deep learning it is none the less desirable (Cartwright, 1997). Performance needs acknowledgment. The university, the student and other end-users of tertiary education need norms against which to measure standards of excellence, improvements in performance and as a benchmark of skills and abilities attained in a given topic (Airasian, 1994; Angelo & Cross, 1993). Apart from producing anxiety, the method of achieving and awarding grades is problematic. There are concerns that many of the tools used to evaluate student performance are subjective (Airasian, 1994; Angelo & Cross, 1993; Cartwright, 1997). Students are at the same time vulnerable to, and active in, the assessment process – with many adept at establishing what their assessors require from them to achieve a passing grade. Rather than actively engaging in the topic, students may respond to assessment by reproducing what they think an assessor wants. Perceptions of such subjectivity in assessment are not conducive to deep learning (Cartwright, 1997).

Evaluation of Group Presentation Assessment

The challenge for assessors in Aust 1004 is to create an assessment method that more effectively evaluates and rewards the type of learning (deep learning) that the topic requires. To this end a number of assessment methods were considered.

Tutor Assessment of Group Presentations

In Aust 1004 students are assessed via participation (10%), two essays (30% each) and a group presentation (30%). While assessment of participation and essays may be open to claims of subjectivity and thus promote surface learning rather than deep learning, the marking criteria are mediating factors (Airasian, 1994; Angelo & Cross, 1993). The greatest concern however, lies with group presentations where the criteria are more flexible. There is the concern that rather than deeply engaging in the topic, students may seek to reduce their anxiety by establishing their tutor's position in relation to their presentation subject and then supporting that position to increase their chances of achieving a good grade. Such reproduction may be self-affirming for the tutor but it is not conducive to student deep learning. While this may be overcome through tutors not expressing their views on topic matters, it does not produce positive learning experiences for students or reduce their anxiety (Cartwright, 1997).

Note: *This reflects my own experiences. In order to retain objectivity I tried not to let students know what my opinions were on the subjects under discussion. After giving a presentation the students told me that they found it frustrating doing a presentation for me as they did not know how to please me. What they wanted to know was my biases, so they could cater to them. If this is what they were aiming for then it was not achieving the deep learning the topic required and it became the catalyst for changing my assessment practises.*

Perceptions of subjectivity in grading could be overcome by implementing an objective form of assessment - a concrete transparent marking criterion for presentations. However, if the criterion is too prescriptive there is the possibility that some students will use it as a checklist for their work that does nothing to remove the problem of perceived subjectivity of the assessor. This may defeat the purpose, which is to promote deep learning in students. One way this may be resolved is by removing the tutor from their key position in grading to one that is less central to the assessment process.

Self Assessment of Group Presentations

Self-assessment is the process whereby the student is involved in the process of developing the criteria against which to measure their performances in a given area. This process differs from self grading (where the evaluation criteria is given rather than self-developed) which can impede deep learning as the student may adopt an instrumental approach and respond to the criteria rather than engaging in the learning process (Cartwright, 1997; Staniforth, 1997). Self assessment assumes that students can take responsibility for setting and judging the standard of their work. It has advantages in promoting deep learning as students can relate their experiences to and of the 'real world' and tailor their learning accordingly. Boud (in Cartwright, 1997) contends that self assessment is a necessary skill required of all students and that it is the foundation of effective learning which makes a significant contribution to a student's professional work after graduation. The process of self assessment can encourage self reflection and help students to develop "...appropriate standards of performance (self-developed criteria) and to apply them to their own work." (Staniforth, 1997:30). Thus self-assessment can be a positive influence in promoting deep learning.

Despite the stated obvious advantages, self-assessment is not without its problems. Among these is that students may adopt an instrumental approach to learning where they seek to satisfy what they perceive to be the requirement of the teaching staff rather than engage deeply in the topic (Cartwright, 1997). If they do this, self-assessment will have no contribution to make in achieving deeper learning. Another problem noted by Brown and Knight (in Cartwright, 1997) is that students may not be competent to judge their own work - self assessment can be affected by overconfidence or by low self esteem. Perhaps the greatest concern however is that researchers have found low agreement between self assessed grades and those awarded by tutors and peers (Staniforth, 1997). However, a student's accuracy in self assessment can improve with experience. Keefer (in Staniforth, 1997) found that as students gained more experience at tertiary study they were able to more accurately self assess their work (though this is mediated by their self concept and past performance as measured by their Grade Point Average). Staniforth (1997) noted that students found self assessment difficult and many thought that it should be the responsibility of the

tutor/lecturer. These problems suggest that self-assessment alone may not be enough to promote the deep learning desired for Aust 1004.

Peer Assessment of Group Presentations

Peer assessment is the process whereby students assume responsibility for providing feedback, comment and assessment of the work of fellow students. As a method of assessment it has advantages (as does self assessment) in that it replicates the professional and working environment (Cartwright, 1997). Professional environments often require the independent judgment of others' work. Peer assessment can help "...students to become more autonomous, responsible and involved." (Cartwright, 1997:57). It can create a more supportive environment where students have the advantage of greater feedback on their performance while at the same time developing an ability to analyse the work of their peers. Peer assessment is useful as it allows students to evaluate their own performance in relation to their peers (norm-referenced assessment) against a negotiated set of criteria (criterion-referenced assessment) (Print, 1993). In addition, peer assessment has the added advantage in that it can reduce the marking load of assessors (Cartwright, 1997).

Peer assessment does present some problems though. As with self-assessment, students may not have the competency or experience to evaluate their work or that of others. The process is open to the criticism that students may not take the process seriously and evaluate performance on the basis of its entertainment value or the popularity of the student being assessed. A further criticism of the process is that students may be wary of the process for fear of bias, discrimination or misunderstanding. Some of those who have implemented peer assessment have found it to undermine student confidence and peer solidarity and to create a state of resentment and hostility (Cartwright, 1997). While there are steps which can be taken to overcome many of these criticisms, they are very much dependent on the dynamics of the particular cohort of students involved in the process and the learning environment (Angelo & Cross, 1993). These problems can cause anxiety and promote surface learning and as such are sufficient to bring into question the effectiveness of peer assessment as a tool to aid deep learning.

Mixed Mode assessment of Group Presentations

Mixed mode assessment is a combination of assessments involving the tutor, the student (self-assessment) and peers (peer assessment). It is a process whereby group presentations are assessed with one third of the grade awarded by the students giving the presentation, one third by the students' peers and the final third by the tutor (each interest in the assessment process being equally represented). This mode of evaluation overcomes many of the impediments that the other forms of evaluation suffer when administered alone. While the tutor has a significant role in the process (the tutor 's grading can prevent students from achieving more than a low credit – 66%) perceptions of tutor bias are lessened, particularly where the criteria for peer and self assessment are the result of negotiation and the tutor input into the final grade is mediated by student assessment. The potential for students to grade work inappropriately through inexperience, low self esteem or bias towards or hostility against their peers is mediated by the marking criteria (against which they can be challenged to justify a given grade) and the grading by the tutor. The potential for popular students to benefit is mediated by the others in the group - grading is not individual but awarded for group performance. Staniforth (1997) found, when evaluating self assessment, that many of his students would have preferred to be peer assessed and there was an underlying feeling that the responsibility for self assessment was an imposition – mixed mode assessment addresses these concerns. Finally the concerns of many academics, that presenters and peers may not take grading seriously and give inflated grades, are addressed (Staniforth, 1997). The inconsistency between self and peer assessment and the assessment that an experienced tutor would award is reduced as the tutor in this process has the power to reduce inappropriate grades significantly (however this power should be exercised with caution otherwise it may engender feelings of powerlessness and reinforce the surface learning strategies which mixed mode assessment is trying to overcome).

From a procedural point of view mixed mode assessment, as used in Aust 1004, appears to incorporate the best features of the three modes of assessment and addresses their attendant shortcomings. But the central question is how do students perceive the process - does the current practice of mixed mode evaluation for group presentations in Aust 1004 promote deep learning and if so is it an improvement (best practice) over other forms of assessment?

Methods and Procedures

In the second semester of 2000 a preliminary evaluation of mixed mode assessment was conducted involving two groups of Aust 1004 students. One group met for tutorials from 4-6pm on Thursdays (n = 16) and the other met from 9-11am on Tuesdays (n = 12). Of those undertaking the topic, approximately fifty percent were doing so because it was a first year course requirement, almost one third were overseas students who were in their later years of study, the remainder were taking the topic as an elective. In their first tutorial after introductions, students were informed of the advantages of mixed mode assessment. The variation between what is required (as per the course Handbook) for regular assessment and the proposed mixed mode assessment was discussed. The major variation being that the 300 word outline, bibliography and evaluation to be submitted by each presenter one week after the presentation instead be submitted to the tutor prior to presentation in order to facilitate immediate feedback (just as one would in real world situations). Students were encouraged to discuss the marking criteria and present a case for its modification or adoption (see Attachment 1 - a criterion based assessment form listing four main areas against which presentations could be evaluated). After discussion students were given the option of choosing between assessment as per the course Handbook or mixed mode assessment – all opted for the latter and accepted the marking criteria without alteration.

The grades that students achieved for their presentations ranged from 76% to 93% with the average grade being 84% (Distinction). At the end of the topic when all students had presented they were asked to evaluate their experience of mixed mode assessment (see Attachment 2 – Survey: Mixed Mode Evaluation of Group Student Presentation in Aust 1004) on a Likert scale from 1 to 5. A number of assumptions underpinned the questions asked. They were: a) that mixed mode assessment encourages deep learning (Q1, Q2, Q5 & Q6); b) that students prefer deep learning (to more fully engage in the topic) (Q1); c) that students would change their learning / presentation style in response to mixed mode assessment (Q2, Q5 & Q6) and d) that those who prefer mixed mode assessment would find the evaluation criteria clear and are happy with their peers' ability to assess them (Q1, Q3 & Q4).

Results and Discussion

The following table shows the means, standard deviations and median scores of 25 students who responded to the six questions of the Survey: Mixed Mode Evaluation of Group Student Presentation in Aust 1004.

Table of Results

	Mean	Std. Deviation	Median
Question 1	1.44	.87	1
Question 2	3.80	1.26	4
Question 3	4.28	.94	4
Question 4	4.44	.92	5
Question 5	4.08	1.15	4
Question 6	3.20	1.66	4

The results showed that students overwhelmingly preferred mixed mode assessment to other forms of assessment (as indicated by responses to question 1) and that it had encouraged them to encompass more and make greater efforts to present their work in a way which was more acceptable to their peers (as indicated by responses to questions 2 and 5). It also indicates that students changed the way they prepared for presentation as a result of the assessment method (as indicated by responses to questions 2 and 6).

There was a significant correlation between question 1 and questions 3 ($r = .465$ $p = 0.019$) and 4 ($r = .828$ $p = 0.000$) and between questions 3 and 4 ($r = .530$ $p= 0.006$) which suggests a significant relationship between assessment criteria, the ability to use that criteria and a preference for mixed mode assessment (see Attachment 3). A relationship was also found between questions 5 and 6 ($r = .428$ $p = 0.033$) indicating that influence of peers (and self) in mixed mode assessment is a significant factor for students when it comes to the choice of presentation style.

In general the results support the assumptions that: a) mixed mode assessment encourages deep learning; b) students prefer deep learning; c) students would change their learning style in response to mixed mode assessment and d) that those who prefer mixed mode assessment would find the evaluation criteria clear and are happy with their peers' ability to assess them (as evidenced in student responses to Q6).

Discussion

It was not possible to undertake quantitative analysis of differences that may result from gender, age, self esteem, or academic experience due to the limited number of students involved in the study. However, there is evidence to support a continuation of mixed mode assessment for group presentations on the basis that it encourages active engagement in the topic – a necessary condition for deep learning. Students appear to be highly satisfied with the process and there is evidence of deep learning as indicated by one student's comment that it encouraged them to "...present in ways that are less conventional, thus challenging me in my presentation style". There was a clear indication that the mixed mode assessment method forced students to consider factors other than using terminology to impress the tutor as indicated by this statement "...although you may understand a technical term or sentence, you express it clearer to ensure everyone understands". Anxiety was reduced (as one student said "It allows you to be more relaxed while presenting. This, in turn, allows you to give a more effective presentation") and there were comments to the effect that presentations were more succinct, fun and stimulating both for the presenter and audience. These factors combine to produce an environment conducive to deep learning (and better grades). At least three presentations were highly polished with one group trialing their presentation at Taoundi College (an Indigenous community college) and two other groups, presenting on Indigenous art, coordinating their presentation with the Flinders Art Museum and using originals from the collection (one student is using what she learned to import Indigenous art into the United States of America). The application of topic and presentation material to real world situations indicates deep learning and reinforces the usefulness of the mixed mode assessment process

In implementing mixed mode assessment there are some areas that need attention. While there is no indication that it affected the outcome, the literature suggests that it is crucial to the assessment process that students negotiate the assessment criteria (Cartwright, 1997; Staniforth, 1997). Since the assessment criteria were prepared in advance and the tutor was in a position of power, as the students had not yet established a working relationship, in essence it was ratified rather than negotiated. Another area of concern is the competency of students to assess. This is less of an issue in this case as Aust 1004 is a second semester topic. Students have had the benefit of matching their grade expectations with what they actually achieved in first semester and thus have some experience of grading expectations. These are important areas of concern because if they are not addressed the literature suggests that it may result in students adopting an instrumental approach to learning rather than engaging in the more desirable deep learning.

Bibliography

Airasian, PW 1994 *Classroom Assessment*, McGraw-Hill, USA

Angelo, TA and Cross, KP 1993 *Classroom Assessment Techniques: A Handbook for College Teachers*, Jossey-Bass Inc, San Francisco, California

Boud, D 1991in Cartwright, N1997 *Assessment and Feedback – a Handbook for Tertiary Teachers*, University of Ballarat.

Brown, S and Knight, P (1994) in Cartwright, N 1997 *Assessment and Feedback – a Handbook for Tertiary Teachers*, University of Ballarat.

Cartwright, N 1997 *Assessment and Feedback – a Handbook for Tertiary Teachers*, University of Ballarat

Gibbs, G 1992 in Cartwright, 1997 *Assessment and Feedback – a Handbook for Tertiary Teachers*, University of Ballarat

Keefer, KE 1971 in Staniforth, M 1997 'Student self-assessment and self-grading in Archaeology' in Silins, H Murray-Harvey, R and Orrell, J (eds.), *Improving University Teaching & Learning*, Flinders University Press.

Print, M 1993 *Curriculum Development and Design*, Allen & Unwin Pty Ltd, Australia.

Staniforth, M 1997 'Student self-assessment and self-grading in Archaeology' in Silins, H Murray-Harvey, R and Orrell, J (eds.), *Improving University Teaching & Learning*, Flinders University Press.

Attachment 1

Marking Criteria: Aust 1004 Group Presentations

(All participants must mark the Tutorial Presentation Criterion Based Assessment Form)

The areas are:

Participation and Involvement of Tute group

(Presenters share equally in the work, Tute group is involved in meaningful and relevant dialogue and interest is maintained)

Content/Clarity

(Well structured and easy to follow, presenters are familiar with the topic and content, make reference to social constructs and show the implications for Indigenous Australians)

Presentation/Delivery

(Presenters work well as a team, they make appropriate use of resources and they manages the tute group)

Presentation Plan

(Evidence of preparation, research and rehearsal)

Note: You are to give a mark out of **2 1/2** for each area for a **maximum** total of 10 marks.

Attachment 2

Survey: Mixed Mode Evaluation of Group Student Presentations in Aust 1004 (1/3 Tutor 1/3 Presenters and 1/3 Peers)

Please circle what you think is the most appropriate response.

Q1. What is your preferred method for evaluation of presentations?

Prefer Mixed Mode (currently used in your tute)		Unsure		Prefer Tutor only
1	2	3	4	5

Q2. Do you think mixed mode evaluation encourages you to encompass more than you would if you were presenting only to your tutor?

No		Unsure		Yes
1	2	3	4	5

Q3. Do you think the criteria for assessing presentations is clear?

No		Unsure		Yes
1	2	3	4	5

Q4. Do you think your peers can properly assess your Presentation?

No		Unsure		Yes
1	2	3	4	5

Q5. Do you feel that peer and self assessment force you to make greater efforts to present information in a way that is more acceptable?

No		Unsure		Yes
1	2	3	4	5

Q6. Does the form of assessment used affect your preparation for Presentations?

No		Unsure		Yes
1	2	3	4	5

If so, how?

.....

Attachment 3

Correlations

		Q1	Q2	Q3	Q4	Q5	Q6
Q1	Pearson Correlation	1.000	-.221	-.465*	-.828**	-369	-.150
	Sig. (2-tailed)	.	.289	.019	.000	.069	.474
	N	25	25	25	25	25	25
Q2	Pearson Correlation	-.221	1.000	.191	.441*	.356	.080
	Sig. (2-tailed)	.289	.	.360	.027	.080	.704
	N	25	25	25	25	25	25
Q3	Pearson Correlation	-.465*	.191	1.000	.530**	.249	.070
	Sig. (2-tailed)	.019	.360	.	.006	.230	.740
	N	25	25	25	25	25	25
Q4	Pearson Correlation	-828**	.441*	.530**	1.000	.439*	.351
	Sig. (2-tailed)	.000	.027	.006	.	.028	.085
	N	25	25	25	25	25	25
Q5	Pearson Correlation	-.369	.356	.249	.439*	1.000	.428*
	Sig. (2-tailed)	.069	.080	.230	.028	.	.033
	N	25	25	25	25	25	25
Q6	Pearson Correlation	-.150	.080	.070	.351	.428*	1.000
	Sig. (2-tailed)	.474	.704	.740	.085	.033	.
	N	25	25	25	25	25	25