Using assessment within course structures to drive student engagement with the learning process

John Mills*, Chris Glover*, Valda Stevens**
* Sheffield Hallam University; ** Open University

Abstract

Assessment has unrivalled potential to draw the student into the learning process. However, this potential may not be met fully if the assessment secures engagement and the feedback does not. In this paper some tactics by which feedback can encourage engagement are described. Group peer assessment and feedback brings students closer to the evaluative processes that underpin learning. In this case the responsibility for grading is an important motivating factor. Conversely, written feedback that focuses on future tasks may be obscured by grades and by comments or model answers that seek to justify these grades. Eliminating grades and reshaping model answers can mitigate this problem. Successful implementations of these approaches are described.

Introduction

Gibbs and Simpson (2004) have identified conditions under which assessment supports student learning. These conditions span the broad themes of engagement and feedback and include the need for feedback to be timely and of high quality and for active student engagement with the feedback.

Investigations carried out at Sheffield Hallam University (SHU) and the Open University (OU) within the framework of the Gibbs and Simpson conditions have prompted changes in the design of course assessment. These changes are intended to provide opportunities for students to engage more fully in evaluating their own progress and therefore in the learning process. More immediate and effective feedback is a key aim.

Although based around common aspirations, these activities employed different tactics and are most easily considered separately. At SHU, the focus was on involving students in the assessment process through a peer assessment activity. At the OU, the accent was on redesigning the feedback provided to students. In this paper, they will be treated sequentially. In each case, the study will require further evaluation.

Group peer assessment of case study reports

Introduction

The coursework for a level 5 microbiology module consisted of five case study reports spaced every two weeks. Working in groups, students produced a group report for each case study, based on six questions. The report was brought to class in draft form. Four members of staff were involved in 30 minute discussions with each group. They asked each student questions on any part of the case study together with peripheral questions to check student learning. Students then modified their group reports and handed them in for marking two days later. Individual student marks were calculated by staff from the group mark which was modified according to student peer assessment ratings for the individual work done. Depending upon staff pressures, the reports could take up to two weeks to mark and return with feedback.
The coursework appeared to meet some of the Gibbs and Simpson conditions. The activities spread the learning load, and linked the feedback to the purpose of the assignment but failed to engage all students in active learning. Students tended to research one part of the case study and not engage with the rest or not engage with staff during class time. Stronger students dominated the sessions. Staff marking failed to provide feedback quickly enough to support the next case study. These defects were reflected in a general levelling of marks across the exercise with no progressive improvements. The exercise was also very expensive on staff time.

In an effort to improve feedback response, peer and self involvement in the assessment process were introduced. Students marked and provided immediate written feedback. The aim was for the assessment itself to become an integral part of learning.

**Introducing peer-assessment.** The number of case studies was reduced from five to four and the number of questions from six to five. The group reports were handed in two days before the class. In class each group marked an anonymised report of another group using a marking plan. All students within each group had to work together to discuss the answers of the unknown group, using the marking plan and their own skills of discrimination in sifting through the information. One member of staff was present (compared to four before) to act as arbiter and give immediate feedback when problems of interpretation arose. Marks were placed on a separate score sheet, and were justified within the student feedback section. After marking, the reports and score sheets were handed in and staff returned the reports to the relevant groups. Students then marked their own report according to the same criteria and arrived at an agreed mark before receiving the score sheet and feedback from the other group. Marks were compared and, if within 15%, the group was allowed the higher mark. If the scores were outside this limit then the member of staff arbitrated.

Individual student scores for the first three case studies were negotiated by students themselves within their groups using three strategies (Race 2001)

- **Case Study 1.** Multiply the group score by the number of students in the group and allow them to distribute the marks between the members.
- **Case Study 2.** Multiply the group score by the number of students in the group and ask them to calculate their own scores according to the way they feel the work was shared, to a maximum differential of 20%.
- **Case Study 3.** Award a mark for the group but ask students to peer assess an additional mark for their own contribution - between one and ten extra marks.

Students were allowed a free choice of which system to use for Case Study 4.

Students were introduced to the pedagogy behind the exercise and trained in marking using a "poor answer" and a "reasonable answer" to an unseen case study. A marking plan was provided and the resultant scores produced were within 15% of each other for each answer for the twelve student groups in this cohort. An interesting outcome of the training session was that students were able to answer a range of questions on the case study at the end of the class without resort to their notes.

For this cohort a member of staff moderated all reports to allay initial student concerns about any perceived unfairness of the peer assessment process. This was completed in two hours per case study compared with two weeks in the previous model.

**Results**

**Student performance.** The majority of students understood and welcomed the approach. However, a few were very concerned about the perceived fairness of the peer marking and the fact that their individual contributions might be degraded by less able students within their group.

The first case study scores showed very little correlation (Table 1) between peer group and own group marks but better correlation between staff and group marks. A serendipitous mistake
allowed students to see the peer group marking and feedback during the second case study before they marked their own reports. The result was almost perfect correlation between the sets of markers. The third case study showed better correlation between peer group and group scores - now mostly within 15% but the fourth case study saw most groups over-marking their own reports with no correlation between any of the markers.

Table 1. Correlation coefficient for sets of markers: G/P own-group/peer group; G/S own-group/staff; P/S peer group/staff.

<table>
<thead>
<tr>
<th></th>
<th>G/P</th>
<th>G/S</th>
<th>P/S</th>
</tr>
</thead>
<tbody>
<tr>
<td>case study 1</td>
<td>0.1173</td>
<td>0.5622</td>
<td>0.4689</td>
</tr>
<tr>
<td>case study 2</td>
<td>0.9302</td>
<td>0.9227</td>
<td>0.8711</td>
</tr>
<tr>
<td>case study 3</td>
<td>0.5256</td>
<td>0.7493</td>
<td>0.8487</td>
</tr>
<tr>
<td>case study 4</td>
<td>0.1238</td>
<td>0.0049</td>
<td>0.0021</td>
</tr>
</tbody>
</table>

Groups worked well together in class and very quickly learnt to penalise students not attending or contributing through peer assessment. The group marks (Figure 1) generally improved over the exercise. The lower marks for four of the groups in Case Study 2 reflected a mistake in diagnosing the disease.

Figure 1. Group scores (%) for each case study

The individual student marks suggested an improvement in the proportion of First class grades achieved by this cohort (Table 2).

<table>
<thead>
<tr>
<th>Range</th>
<th>cohort 2004</th>
<th>cohort 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>48%</td>
<td>72%</td>
</tr>
<tr>
<td>Upper second</td>
<td>40%</td>
<td>17%</td>
</tr>
<tr>
<td>Lower second</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>Third</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fail</td>
<td>9%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Table 2. Individual student scores over two years
**Student attitudes.** After the fourth case study, students were asked to complete a questionnaire covering their attitudes to the exercise (Table 3). A four point scale spanning from highly positive to highly negative was used. Free format responses were also encouraged. Overall responses were favourable. A significant issue reported was in marking overlong answers in which markers had to search for salient points. Students will be shown examples of "concise" and "overlong" answers in the next practice session.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Positive Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 Did you find PA useful?</td>
<td>80.5 %</td>
</tr>
<tr>
<td>Q2 Were the marking criteria helpful?</td>
<td>75.6 %</td>
</tr>
<tr>
<td>Q3 Did you have enough guidance about PA?</td>
<td>87.8 %</td>
</tr>
<tr>
<td>Q4 Did PA help you understand your own work better?</td>
<td>78.0%</td>
</tr>
<tr>
<td>Q5 Did assessing each other’s work help you to see where your own work might be improved?</td>
<td>82.9%</td>
</tr>
<tr>
<td>Q6 How did you feel about others assessing your work?</td>
<td>85.4%</td>
</tr>
<tr>
<td>Q7 How did you feel about assessing other’s work?</td>
<td>87.8%</td>
</tr>
<tr>
<td>Q8 Were you happy with the way your work was assessed?</td>
<td>82.9%</td>
</tr>
<tr>
<td>Q9 Was the feedback received helpful?</td>
<td>75.6%</td>
</tr>
<tr>
<td>Q10 Was the feedback fair?</td>
<td>92.7%</td>
</tr>
<tr>
<td>Q11 Do you think the feedback you gave was helpful?</td>
<td>75.6%</td>
</tr>
<tr>
<td>Q12 Do you think the feedback you gave was fair?</td>
<td>95.1%</td>
</tr>
</tbody>
</table>

Table 3. Percentage positive responses to questions on the student feedback questionnaire (n = 54 students, 76% response rate).

Overall satisfaction for each group was computed using a four point numerical scale and averages across all questions. Ten groups returned scores of greater than 80% of the maximum 48. Two groups returned 70% and 54% respectively. These groups contained the students who were most openly opposed to the process from the start and who maintained their opposition throughout.

The free format responses provided useful evidence of the links established with learning. They included the following positive and negative comments.

...highlighted points we missed
I was able to learn how to improve for the next case study
showed where marks were gained and lost
...good to see how answers should be structured
...it helped me establish what was a “good” answer
...gave immediate feedback
...allowed us to see what was needed to achieve high marks
...showed what the question was really asking
...it made me realise I need to give more concise and relevant answers
I learned through marking
... peer comments were constructive
...the marks were normally justifiable
I don't like the responsibility and don't feel I have enough knowledge or experience. I don't think I can mark people's work fairly; marks should be a lecturer's judgement, not mine. I still don't like peer assessment and think it's a waste of time! I expect my work to be marked by a lecturer – we could have spent the large amount of time we spent doing this doing proper lectures! We are paying £1100 tuition fees and from this course we have got nothing for our money. We have marked our own work!

Discussion

Overall, the outcome of the exercise was successful in terms of engaging students with the learning process. Responsibility for the reports, marking and feedback passed to the students themselves, who, after a shaky start, engaged well with the process. The correlation between own group and peer group marks was a cause for concern. Since correlation between marks was far better when groups received the marks and feedback from the peer group, it may be the better system to use since students can still discuss and negotiate around the marks given by the peer group with or without staff arbitration and this may temper own-group over-marking. It is still important, however, that groups mark a peer group report before marking their own.

Student feedback was most encouraging with most concerns raised about the marking plans. Suggestions were made to simplify the mark plan, including providing bullet point information.

Class observation showed excellent student involvement, enthusiasm and rigour. Two students failed this part of the coursework through non-involvement in preparing two of the case studies: they were given zero marks by the rest of the group. Several students were marked down by their group for non-attendance of the class session even though they had contributed to the group report. Initial concerns about fairness of the system were soon allayed but there were still a few students who preferred to passively receive information and not take an active part in the learning process. Habeshaw (1999) has suggested the responsibility of the teacher in this situation, "what they (students) pay for is the opportunity to have an education.... our job becomes that of creating the highest quality educational environment, and of motivating the students by initiating a range of learning experiences, so as to make it possible for them to qualify well so long as they make their contributions to it all." The negative perceptions will be used in next year's training session and may trigger a different response.

Redesign of Feedback

Introduction

OU students complete Tutor Marked Assignments (TMAs) which they send to their tutors for grading and feedback. Tutor Notes include detailed marking schemes which support this grading and feedback. Tutors give written feedback comments both on the student script and on a summary cover sheet. In addition, Student Notes with summary answers for all, or selected parts, of the assignment may be prepared to accompany the returned script and tutors may highlight and/or annotate particular sections to tailor these to individual need.

Whilst students are happy with the amount and detail of feedback they receive, questionnaire and telephone interview investigations indicated that they made little use of the feedback in feeding forward to future work (Brown et al, this volume). A number of changes, therefore, have been made to the assessment processes with the aim of providing feedback that feeds forward to students' future work and thereby secures greater engagement. A range of tactics have been employed on a number of Science courses. Here the focus is on just two tactics– the removal of numerical grades and the provision of enhanced Student Notes. In both cases, the changes are limited in the context of the overall learning experience and the evidence of benefit is similarly tenuous at a statistical level. The comments made here will rely more on professional judgement.
Removing the numerical grade

The first assignment of the level 4 module Discovering Science is formative in the sense that it does not carry a summative grade. However tutor input has been standard in that written feedback and marks have been supplied. In order to encourage students to engage in reading and responding to the guidance provided by their tutor, particularly in the development of appropriate skills, the process was changed by removing the allocation of the numerical grade such that all feedback was through tutor comments on the script and cover sheet.

Limited samples of tutors and students were surveyed to assess their attitudes to this reform. Response rates were too low to be statistically meaningful. However, some indications could be extracted.

21 tutors indicated that effective feedback could be given without recourse to quantitative indications. Their comments included:

- **I felt more focused on the feedback because I didn’t need to worry about assigning marks.**
- **It was easier to focus on areas for improvement.**
- **Without the mark the students read and act upon the feedback in a better way.**

There was less consensus on the consequent influence on student learning though one tutor commented:

- **I am getting slightly more comments from students in which they acknowledge the feedback.**

Over 80% of the student respondents said that they did take time (average 15-30 minutes) to look at the feedback from their tutor. Many found the feedback comments “very useful” or “invaluable” in tackling the next two TMAs. Comments included:

- **It helped me to understand what was required and how to set things out.**
- **They gave me confidence as my tutor highlighted where I had got things right so I could build on these skills in future TMAs.**

Views on the loss of a grade were mixed.

- **It is nice to know how well you scored - more of a morale booster.**
- **It actually took the pressure off.**
- **I could use the first TMA to ‘experiment’, knowing it carried no marks.**

Analysis of script and cover sheet comments using a coding system for written feedback to students (Brown et al, 2003) revealed one interesting difference from previous practice. The cover sheet summary for the assignment carrying no grade provided a higher level of skills and specific feed-forward comments than were found on other assignments for this course (45% compared with 34% and 30% for TMA02 and TMA07 respectively).

Making skills explicit

The team supporting the level 5 module Biology, Uniformity and Diversity provides Student Notes which accompany the marked TMA, the comments on the script and the cover sheet summary. These Student Notes consisted essentially of model answers. In this form they were likely to encourage retrospective analysis of the assignment rather than future engagement. They have been enhanced to encourage students to focus on learning rather than on marks by explicit linkage between skills required in the present work and the demands of future assignments and the examination. Tutor Notes make clearer the skills for which marks are to be awarded.
The effectiveness of this intervention has been assessed via questionnaire and a script coding analysis. An Assessment Experience Questionnaire AEQ described in Brown et al, (2003) was administered after the end of the 2004 course presentation following implementation of the changes and comparisons made with the results, pre-change. 42 and 94 questionnaires were returned respectively. An increased score (on a 5 point scale) indicates improvement in meeting the conditions for assessment supporting learning. The largest changes were recorded in response to the following statements.

- The feedback shows me how to do better next time (3.76 to 4.41).
- The feedback helps me understand things better (3.77 to 4.02).
- I read the feedback carefully and try and understand what the feedback is saying (4.02 to 4.21).

For those items with a reverse scale, the largest changes occurred for

- I can seldom see from the feedback what I need to do to improve (2.32 to 1.93).
- The feedback does not help me with subsequent assignments (2.51 to 2.07).
- I do not use the feedback for revising (3.01 to 2.57).
- I only tend to read the marks (1.89 to 1.62).

There is a clear pattern of greater student engagement with their feedback after the change. To what extent the enhancement of Student Notes and explicit skills guidance in Tutor Notes contributes to this is unclear.

Analysis of tutor feedback comments using the comment coding system revealed a higher level of feed-forward comments in the script feedback for assignments after the change. Cover sheet feedback was similar in both years.

Student interviews suggest that the Student Notes, being 'pithy', are of use for revision and also for skills development. More students responded that the comments fed forward from one TMA to the next.

Enhancement of feed-forward comments is revealed by both measures and in association with the AEQ data, increased engagement with the feedback to support learning is suggested.

Discussion

It has already been noted that the changes to the feedback and the evaluation are both limited. Although the indications are positive, OU students tend to be highly motivated and may be particularly adept in gaining benefit from pedagogical change that is focused on increased engagement. It is possible that the limited changes made serve the purpose, for both tutor and student, of drawing attention to the feed-forward potential of the tutor comments and the Student Notes. However, no specific conclusions can be drawn about the effectiveness of these tactics until a fuller evaluation is carried out.

Conclusion

The Gibbs and Simpson framework, in common with other analyses, highlights the central roles of engagement and feedback in driving learning. These issues are not separable as the feedback processes must be designed to secure engagement. This paper has discussed two tactics for securing such engagement.

Group peer assessment is a strong intervention with potential for a major shift in performance but with attendant difficulties in implementation. It affects students and therefore they may react strongly. Reshaping of feedback is a less radical intervention, although the removal of grading, even on an assignment with zero summative weighting, confronts the grade centred culture within which universities work.
For the studies reported here, there is evidence of benefit. The Gibbs and Simpson conditions have been met more fully. Engagement has been increased by refocusing the source or the style of the feedback offered to students.

References


Brown, E and Glover, C. this volume Refocusing written feedback

