This Teaching Cycle of the PGCTHE portfolio has been submitted to CADAIR with the permission of the author detailed above. It is to be used as a resource for future PGCTHE candidates and other staff as part of their professional development at Aberystwyth University. If you wish to cite this work then please contact the author. Contact details can be found at http://www.aber.ac.uk/en/directory/.
CHAPTER 4

Teaching Development Cycle 3

PEER ASSESSMENT OF SEMINARS
4.1 Introduction

The evaluation of one’s own work is essential for a successful professional career, and yet some suggest that it is a skill that many university schemes often fail to cultivate (Biggs 1999). One method thought to develop these ‘internalised standards of competence’ (Biggs 1999), and thereby enable reflective thinking and self-direction, is self- and peer-assessment (Boud, 1995; Strachan and Cox, 1996; Gibbs 1998). That is formalising the grading of student work by the student or by his or her contemporaries. There are numerous theoretical advantages to self and peer assessment:

i. Self/peer assessment allows the student to develop a sense of what is and what is not acceptable for a subject and expedites the process of becoming an autonomous learner (Biggs 1999)

ii. Gibbs (1998) claims that the comments on peer-marked work are “socially amplified”, that is have a much greater effect with respect to the original student than comments from a tutor they hardly know.

iii. Students can receive feedback on their own work quickly (because of the increased manpower, rather than weeks later when a tutor has had a chance to mark), enabling students to relate comments given to the process of completing the work (i.e. before they have forgotten).

One common criticism of peer assessment is that the feedback is often inexpert, unfair, or focuses on the wrong issues relative to the assessment set. Clearly there is a need for student comments to be moderated by those of a tutor, so that an individual student is not given the wrong impression or common misconceptions are not allowed to persist. However, Gibbs (1998) suggests that inexpert peer feedback hardly matters so long as the assessing student is directing efforts to appropriate learning activities. That is, peer assessment of work is possibly best seen not as a means for assessment of that work per se, rather a device for getting students to apply themselves to learn.

In the spring of 2001 the UK was subject to an outbreak of foot and mouth disease (FMD), which caused the cancellation of all farm visits and some other activities by students of IRS as a precaution against the further spread of the disease. On two such afternoons for RD15010 I decided to run a series of student led seminars. Students were given a title and were then asked to present a seminar followed by a few questions. All students were asked to assess all of the seminars, other than their own,
according to given criteria. Finally, students were asked to give feedback on how they viewed the exercise.

4.1 Methods

The ‘RD15010: Animal Science’ module in 2001 was a class of 15 students in the first year of a HND in Agriculture. All of the students had strong agricultural backgrounds but generally had entered IRS with little knowledge of, and few qualifications in the traditional sciences. Briefly, the objectives of the module were to provide an underpinning scientific knowledge appropriate to an understanding of effective livestock production. In particular, the module covered anatomy, physiology, reproduction, endocrinology and issues concerned with the health welfare in domestic farm species. This module was worth 10 credits at level one (22 lectures plus 4 practical sessions/farm visits) and was taught between three people including myself. My input into the module was to teach the anatomy and physiology of farm species (cow, sheep, pig and chicken), with particular emphasis on the digestive tract, the mammary gland and muscle physiology. I was allocated a total of 8 lectures and 2 of the practical sessions in order to do this. The module description and the initially planned scheme of work are given in Appendix 4.1.

The outbreak of FMD in the United Kingdom caused the cancellation of all visits by our students to places off the IRS campus. The consequence of this for this group of students was huge, as many modules, including RD15010, had to cancel planned events. I had intended to have a visit for one afternoon and to bring in from a slaughterhouse some sheep and pig alimentary tracts (guts) for dissection for the other. The latter event also had to be cancelled because at this time no livestock were going through slaughterhouses. In this context, and to prevent the students having yet another afternoon of no planned activity, I decided to hold two afternoons of seminars. Students were expected to do some research and then present a 10-minute seminar followed by 2 or 3 minutes of questions on what they had found. It was probable that these students had not previously had to present information to their peers before, and so I took pains to ensure that the experience for them was not too taxing. The 15 students were presented with a list of 30 seminar titles (see Appendix 4.2) two weeks before the first seminar session, from which they were asked to choose one. The titles were based either directly on what I knew their course content to be, or on information that was easily available in the library or on the Internet. Three students did not turn up for the seminars, and when questioned later said that they were ill, and thus only 12 seminars
were presented. Students were given help in locating information if they requested it. Students were instructed that they could cover any aspect of a title if they found it too large to cover in 10 minutes, but that they had to speak for 10 minutes. The seminar titles and the date that they were given were are shown below:

**Session One (held on 16 March 2001)**
1. Foot and mouth disease (FMD)
2. Animal stress and meat quality – *Not held - student ill*
3. Manipulation of fertility in farm animals
4. Milk quality and milk pricing
5. Characteristics of good stockmanship
6. Heifer rearing
7. Measuring dairy herd performance

**Session Two (held on 23 March 2001)**
8. What is ‘milk let-down’, and how does it work?
9. Differences between the 1967 and 2001 FMD outbreaks
10. The welfare aspects of cubicles for housing dairy cows
11. The increase in the average milk yield of dairy cows. *Not held-student ill*
12. Bovine spongiform encephalopathy (BSE)
13. Increased role of supermarkets on farm incomes. *Not held - student ill*
14. Can mastitis ever be eradicated?
15. Should Europe allow the use of the hormone *bovine somatotropin* to increase milk yields?

On the day of the seminars, students were asked to rate the quality of their classmates’ presentations as a score out of 10 for each of the following five criteria. Before the first session, and to a lesser extent before the second, I elaborated on each of the five criteria to make the meanings of the principal words clear. I did not, however, instruct the students on how many marks to give in any particular instance. The criteria were:

1. **Structure of the presentation**: Clear introduction of main points; appropriate ordering of material; appropriate diagrams and tables; clear summary of main points at end; effective ending.
2. **Delivery**: Appropriate speed; audible; emphasis of key points; enthusiastic; not reading from a script!
3. **Interaction**: Engaging stance, gestures and eye contact; some but not too much movement; friendly demeanour; facing or hiding?
4. **Visual aids**: Clarity; relevant content; appropriate timing of use; not over reliant on visual aids.
5. **Dealing with questions**: Sound knowledge; relevant answers; confident responses

The marks awarded by the students, as well as the marks that I awarded were collated and the following week passed back to the student anonymous of which student had
awarded which mark, along with some written comments by me. These marks did not contribute in any way to the assignment marks for this module, and in this sense this exercise was entirely a formative assessment. No action was taken against those students that did not present, nor were they asked to present at another time.

A week after the seminars were completed the students were given a sheet of paper with the following questions on it (student feedback was returned anonymously):

1. What did you learn from giving a seminar and assessing your classmates?
2. What would you have done differently?

4.3 Results and Discussion

The average mark awarded by the students for all of the assessment criteria over all of the seminars was 5.05 ± 0.19 and that for the tutor (me) was 4.73 ± 0.21 (average ± standard error). The students on average tended to be slightly more generous (P<0.1; Student’s paired t-test) than I was over all criteria. The average student score for each of the criteria and the average score awarded by me is shown in Table 4.1. When each of the criteria were taken individually, only the scores for ‘Dealing with Questions’ were different between the students and the tutor (P<0.05; Table 4.1). No one criteria scored significantly worse than any of the others, either for the student or tutor marks. The marks given for each of the criteria are shown in full in Appendix 4.3.

<table>
<thead>
<tr>
<th>Assessment Criteria</th>
<th>Average Student Score</th>
<th>Average Tutor Score</th>
<th>Student vs. Tutor P (SED)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>5.08</td>
<td>4.58</td>
<td>NSa (0.359)</td>
</tr>
<tr>
<td>Delivery</td>
<td>4.76</td>
<td>4.42</td>
<td>NS (0.207)</td>
</tr>
<tr>
<td>Interaction</td>
<td>4.66</td>
<td>4.67</td>
<td>NS (0.218)</td>
</tr>
<tr>
<td>Visual aids</td>
<td>5.44</td>
<td>5.08</td>
<td>NS (0.370)</td>
</tr>
<tr>
<td>Questions</td>
<td>5.32</td>
<td>4.92</td>
<td>P&lt;0.05 (0.173)</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>5.05</strong></td>
<td><strong>4.73</strong></td>
<td><strong>P&lt;0.1</strong> (0.141)</td>
</tr>
</tbody>
</table>

aN: Not significantly different (P>0.1); bSED: Standard error of the difference between means

A plot of the average student score for each of the criteria for each of the seminars against the respective score awarded by me (Figure 4.1) showed fairly good agreement between the tutor and student. However, a linear fit to the data was not quite as good as the logarithmic fit shown in Figure 4.1. A 1:1 line (also shown in Figure 4.1) marks the line had the tutor and student marks agreed exactly. The student marks tended to be above this 1:1 line at the lower end of the range of marks, while they tended to be below the line at the higher end of the marks. That is, there is an indication that the
students tended to be more generous at the lower end and less generous at the upper end, and therefore that I used a greater range of the marks than the students did. It is accepted however that there are too few students in this study, and the effect too weak, for this supposition to be asserted with confidence. Nevertheless, care does need to be taken with peer assessment if it is generally the case that students are unwilling to award marks at the top and bottom end of the scale.

![Figure 4.1](image)

**Figure 4.1** Correlation between the average mark awarded by students for each criterion for seminar presentations against the mark awarded by the tutor (author).

The students apparently enjoyed the novel experience of grading their peers, and certainly they appeared to engage with each seminar in their efforts to grade each of the criteria. It is perhaps unsurprising that the students chose the more applied of the seminar titles available to them (appendix 4.2), given the vocational nature of the scheme that they were on. Some students asked questions in the query period after the seminars, though student questioning was usually of a basic (structural) nature: the questioning was generally left to me. It is likely that an unwillingness to be seen putting their peers ‘on the spot’ prevented more serious questioning from the students.

It is impossible to know objectively whether or not the students benefited in terms of what they learned from this exercise, and the evidence there is, is equivocal. There was no improvement in the average module mark for this cohort relative to that of the year before or the year after, when seminars of this nature were not held. The average module mark for 2000-2001 (43.9 ± 2.3) was somewhat less than that of 1999-2000 (49 ± 2.0) or that of 2001-2002 (45.2 ± 4.2). Again, it must be stated however, that
comparisons of this nature are flawed because of the great number of uncontrolled variables between years (discussed in Chapter 3).

Biggs (1999) notes that student led seminars can become a surface approach to teaching for undergraduate students (and these HND students were not even that). Often, the only beneficiary is the presenter, and then, only with respect to the topic being presented, the audience “getting yet another lecture given by someone with even more hazardous lecturing skills [than the tutor]...” (Biggs 1999). Certainly, some of the seminars presented as part of this teaching development cycle were less than satisfactory (as evidenced by some poor marks) whilst others were excellent. However, I felt that the process of peer assessment did challenge the students to pay attention and to evaluate critically the information presented. I felt that the revision that these seminars caused the students to do, did increase their understanding for the material. It is impossible to prove this objectively though. The participation rate for the seminars was high despite the fact that that this work did not contribute to the module assessment. I did not, however, explicitly state to the students that this was the case before the seminars had been presented. The lack of visits due to FMD during the semester possibly increased these students’ willingness to participate, as the novelty of ‘afternoons off’ had worn off by the time the seminars were held.

In response to the question “What did you learn from giving a seminar and assessing your classmates?” virtually all students (9 of 11 respondents) said words to the effect that they appreciated the chance to practice speaking in front of their colleagues, and that it was a valuable and enjoyable experience, even if they did not enjoy the presentation at the time. In response to the question “What would you have done differently?” some students thought that my questioning was a little harsh (though I did not think it was at the time; this is something that I will pay more attention to next time), and two students would have preferred not to go through the exercise at all. One student thought it was unfair that I was getting them to assess the presentations. A full summary of the student responses to the questions is given in Appendix 4.4. In my view, the weakest aspect of the presentations was the delivery and interaction of the student with the audience. Training in this and in visual aid preparation beforehand may have helped; on the feedback forms, two students also thought that this would have helped. Another possible refinement might be to allow the students to choose their own assessment criteria before any seminars were given, further involving them in the process. Alternatively, they could have been given a free range regarding the subject that they spoke on, though on this occasion I considered that it would be easier for the
student if they were given a definite question. It was enjoyable seeing the students think about how to assess the seminars; I thought that they did so with considerable maturity given their previous lack of experience.

Peer assessment is a method that I will use again, possibly with other methods of assessment, though clear instruction needs to be given so that the full range of marks is awarded. These seminars were something of a last minute response to changes imposed by FMD and I am absolutely certain that doing them was better than giving the students more afternoons off. I have not repeated these sessions for this module in later years however, simply because there is not enough timetabled contact time in which to do it. Additionally, because farm visits etc. for these students have now resumed, it is possible that participation would reduce unless it was incorporated into the assessment for the module.
CONCLUSION

This portfolio enlarges upon three attempts (amongst a good number of other undocumented efforts) that I have made to improve the teaching that I do. At the beginning, one of the main problems I encountered was facing classes with little in the way of teaching experience and sinking more often than swimming. In the context of higher education and my own experience, another problem for me was the highly vocational nature of the schemes at IRS and the students that enter it. However, whilst trying to maintain a good humour, and while making many mistakes, my teaching has undoubtedly improved over the two and a half years that I have been here. Personal ‘reflection’ is a fashionable teaching ‘buzz’ word which through overuse has probably lost some of its meaning. However, by forcing myself to think about how a particular teaching session has gone and how it could have gone better, improvement has been effected. Undoubtedly, I have found personal reflection worthwhile. The most satisfying aspect of this work has been when I have brought the students to think about the lecture material, and have seen the ‘scales lifting from their eyes’; this was particularly the case for the end of lecture questioning that was detailed in the second teaching cycle. At the very least, I hope to continue to reflect on my teaching, and to look to new resources for new ideas on how I might do it better.
LITERATURE CITED


course in microclimatology. *Journal of Geography in Higher Education* 20
343-353.