1

The effect of marking criteria and exemplars on students' learning during peer- and self-assessment of scientific posters

STEPHEN MERRY & PAUL ORSMOND

BACKGROUND AND RATIONALE

The authors of this report are practicing lecturers with an interest in the influence that assessment practices have on the way that students learn. This "case study" is the combination of four studies investigating student and tutor perceptions of poster marking criteria at Level 1 undergraduate modules within the general field of Biology. Self- and peer-assessment exercises of the students' completed posters together with organised, but informal, formative feedback sessions were used to provide data concerning students' and tutors' perceptions of marking criteria.

The precise learning outcomes of the component individual studies differed, but overall it can be stated that, *'at the end of their participation students should be able to'*:

- explain the meaning of specific marking criteria in a professional biological context;
- provide appropriate formative feedback to colleagues concerning their performance;
- engage meaningfully in the process of peer review as used by professional biologists;
- reflect more on the assessment process as part of their own learning and thereby enhance their learning.

'HOW TO DO IT'

The formats of the four studies were similar, but not identical. The approach described below is a composite which reflects how we would now run such a study.

Stage 1
(4–6 weeks before the poster assessment exercise)
Students were informed that:

- they were required to make a scientific poster, the date of the poster assessment exercise and the topic of the poster;
- posters are a recognised format in which scientific researchers present their results i.e. the assessment was relevant;
- they were required to supply particular materials (i.e. paper headings, adhesives etc) in order to construct their poster and the size of the poster boards;
- they would work in groups of approximately five to discuss the marking criteria, but they would be required to produce individual posters;
- more details would be provided at a later date.

Stage 2
(3–4 weeks before the poster assessment exercise)
Either students were informed what the marking criteria were (if they were tutor provided) and were then allowed time (approximately 30 minutes) to discuss the meaning of the marking criteria in their groups with tutors circulating among the groups to contribute to the discussions.

Or students were allowed time (approximately 45 minutes) to work in their groups to both generate, discuss and refine their own poster marking criteria and agree them with tutors circulating among the groups during this process.

Students were then informed that:

- they would be required to peer-assess the posters of others and self-assess their own posters using the marking criteria they had just discussed and that tutors would also assess the posters using the same marking criteria;
- anonymous peer review was a process utilised by professional scientists which was fundamental to establishing the credibility of scientific publications i.e. they were engaging in genuine professional practice;
- they should regard the self- and peer-assessment activity as a vehicle for developing

specific skills such as self reflection and objective judgement required by professional biologists;

- their posters were to be presented anonymously to reduce any bias in the assessment;
- their self- and peer-assessment would contribute to the overall grades awarded for the exercise and that tutors were interested in the quality of comments made by students in addition to their accuracy of their marking compared to that of tutors;
- more information would be provided at a later date.

Stage 3
(1–2 weeks before the poster assessment exercise)

- Students were given written information concerning the meaning of the individual marking criteria. This information was influenced by the discussions that had previously taken place between tutors and students. It should point out both the meaning of the criteria and the misconceptions which some students seemed to have.
- Students were also given a copy of the marking form they would be required to use for the self- and peer-assessment of their posters and its use was discussed with them, paying particular attention to they types and usefulness of the feedback comments they might provide to their peers.
- Students were given the opportunity to view exemplar posters and to discuss them in their groups and with circulating tutors. They should decide what feedback they would give to the author of the poster and what grades they would award for each marking criterion.
- If student-derived criteria are being used to mark the posters, students should be given, in discussion with tutors, the opportunity to refine the criteria; although such changes should be agreed and discussed with the whole class since changes to the written information provided may be required.
- Students were reminded of the date and time of the poster construction and assessment exercise together with the materials they would need to bring to the session in order to construct their poster, and the time they have available (i.e. 30 minutes) for poster construction.

Stage 4
(The poster assessment exercise)
In an initial plenary session students were informed:

- they will be randomly allocated to two rooms;
- they will be given coloured stickers to attach to their posters;
- they will be given 30 minutes to construct their posters before the start of the self- and peer-assessment exercise;
- they will be required to self-assess their own posters and then move to the other room where they are required to peer-assess all the posters having the same coloured sticker as their own;
- their self- and peer-assessment marking should be independent, i.e. they should not discuss their marks and comments with other students; although tutors were available to provide guidance regarding the usage of the marking forms;
- tutors were interested in the quality of feedback comments as much as the grades awarded;
- in their poster construction they should ensure that only their student number appears on the poster NOT their name.

At the end of the plenary session students were allocated to their rooms and given their individual coloured stickers to attach to their completed posters.

- The format was as described in the introductory plenary;
- Packs of marking forms (one form for each poster to be assessed) were made available to students at the commencement of the self- and peer-assessment exercise;
- At the end of the session tutors collected in the completed marking forms and elicited any informal feedback on their experience of the assessment process from individual students to enable the procedures to be refined for subsequent cohorts.

ADVICE ON USING THIS APPROACH

Tutor discussion with students is the key to the success of the exercise. Tutor discussion should provide to students a) feedback regarding their interpretation and use of marking criteria and b) reassurance that they do have the ability to judge the scientific quality of a poster without the background knowledge necessary to judge the accuracy of the factual detail provided within it.

Posters produced by previous cohorts of students are a good source or exemplars. Students should be given the opportunity to view posters of differing styles and quality. It can help some students realise that attractive posters may, in some cases, have poor scientific depth.

If a sequential allocation of students to rooms and to peer-assessment groups (i.e. a sequential allocation of different coloured stickers) is adopted and these are allocated to students in turn as they leave the plenary session this helps ensure that friendship groups (who are likely to leave together) are separated during the assessment process.

In a 1-hour self- and peer-assessment exercise students should be asked to grade and provide feedback on a maximum of five posters; including their own. This is to enable them sufficient time to write meaningful feedback comments. The number of peer-assessment groups, i.e. the number of different coloured stickers used, should reflect this.

Self- and peer-assessment are skills that improve with practice. Furthermore, informal feedback from students has indicated that they are more willing to engage with the process at a deeper level if they meet it on a second occasion. If possible, repeated self- and peer-assessment exercises should be built into the curriculum rather than being delivered as single isolated events.

TROUBLESHOOTING

In some instances students were unwilling to award low grades, even if they thought they were deserved, to posters that they recognised as being those of their close colleagues. Furthermore, in discussions approximately one month after the completion of the exercise individual students have informed tutors that they felt other students (to whom they did award a low grade) were behaving differently towards them. To address this, tutors need to ensure posters are as anonymous as possible and that friendship groups are distributed into different peer-assessment groups.

Students do find the process challenging. This can be a surprise if they have initially a superficial view of presenting information in poster format. Furthermore, some students feel they are not able to judge the scientific merit of posters whose topic is not strictly the same as theirs. They need reassurance and guidance that they can judge the scientific merit of a poster without being able to judge the accuracy of the detail present in the content.

Finally some students can treat the process in a rather cavalier fashion. Emphasising at the start that they are engaging in a practice which is employed by professional biologists to establish the credibility of their work helps to dispel this.

DOES IT WORK?

Both formal questionnaires and informal feedback from students have indicated that self- and peer-assessment exercises caused students to reflect more on the marking criteria and their learning (Orsmond *et al.,* 1996; Orsmond *et al.,* 1997; Orsmond *et al.,* 2000; Orsmond *et al.,* 2002; Orsmond *et al.,* 2004). To this extent the approach does "work".

The ability of students to mark in an identical fashion to tutors should not be the sole criteria of success of self- and peer-assessment , but it can provide information as to the nature of the learning that is taking place. Our initial studies (Orsmond *et al.,* 1996; Orsmond *et al.,* 1997) demonstrated an overall agreement between student and tutor grades ($r^2 = 0.7$) comparable to that of other studies (Hughes and Large, 1993 and Stefani, 1992) with the agreement being greater for peer-assessment than for self-assessment. Consideration of the overall mark does, however, mask variations between tutor and student with regard to individual marking criteria. For example, students over-marked, compared to tutors for the criteria "visually effective" and "helpful level of detail", but under-marked for the criterion "clear and justified conclusion". The implication is that some students had written a clear and justified conclusion, but did not realise that they had done so. The necessity for dialogue with students concerning individual criteria was shown by these findings.

An interesting, and unexpected, outcome was that our studies, based on comparison of tutor and student grades, have indicated that the nature of the learning that has taken place differs dependent on whether the marking criteria are tutor-provided or student-derived (Orsmond *et al.,* 2000). The use of student-derived criteria might be expected to circumvent discrepancies between tutors' and students' marks for individual criteria since tutors, with their greater experience of interpreting marking criteria, would be expected to more readily understand student-derived marking criteria than, possibly, students understand marking criteria

provided by tutors. In our hands the outcome of using student-derived marking criteria was that students, although having ownership of marking criteria they had constructed themselves, were less able to discriminate between their own individual marking criteria than between those provided by tutors. Student groups either over-marked or under-marked all their marking criteria compared to tutors such that overall agreement between students' and tutors' marks was not enhanced. It may be that the act of constructing their own marking criteria caused students to view their posters in a more holistic fashion. An alternative interpretation of the finding is that students were able to interpret their marking criteria, but had a poor conception of the subject standards, i.e. both students and tutors knew what, for example, the marking criterion "self-explanatory" meant, but, despite the dialogue, they retained different conceptions of how self-explanatory the poster should be to achieve a particular grade.

Our final published study (Orsmond *et al.,* 2002) indicated that the use of exemplars was able to largely overcome discrepancies between grades awarded by students and tutors for student-derived marking criteria. The exemplars were posters produced by a previous cohort of students and served as a focus for discussion and application of the marking criteria. In addition to improving accuracy of marking for individual criteria, feedback from students indicated that the use of exemplars can help students' learning such that higher quality learning outcomes, including reflection, are achieved; although exemplars may not necessarily help students in the process of poster construction. A recent study (Orsmond *et al.,* 2004) has revealed that peer-assessing students were less able than tutors to write constructive feedback comments to the poster authors. Students' feedback comments concerned primarily the quality of the presentation of material with little actual mention as to whether the discussed marking criteria had been met. Tutors' comments, alternatively, concerned primarily the nature and use of the scientific content of the poster in the context of the marking criteria. A possible explanation for this is that students may focus, when constructing a poster, on the poster itself (i.e. the product of their work) whereas tutors may regard the poster simply as a means to enable students to demonstrate the understanding of science they have developed (i.e. to show the process that they have developed).

In summary, the strengths of the approach are that it causes students to reflect more on their work and their learning, but for this to happen, careful planning is required together with the allocation of class time for the activities.

FURTHER DEVELOPMENTS

The authors are currently investigating: How students' perceptions of marking criteria change during the course of the six week poster design and construction exercise; the type of distractions (i.e. student self-derived individual criteria that are distinct from the agreed marking criteria) which influence students' poster design and construction as well as how students use the feedback provided by tutors to enhance their learning.

2

Online calibrated peer-assessment — student learning by marking assignments

VICTOR KURI

BACKGROUND AND RATIONALE

One of the first things that markers have to do to evaluate work is to agree a set of standards. Somehow lecturers have to develop an understanding of what is a good assignment, what is average and what is poor. When I collect a pile of assignments for marking in an area that I have not set before, the first thing that I do is to try to find some of the extremes, and 'calibrate' my marking scheme.

The use of 'calibrated' exemplars can help students become competent at peer review and understand what makes a good (and bad) assignment. I use the web-based Calibrated Peer Review (CPR) system which was developed on a science-based model of peer review (http://cpr.molsci.ucla.edu/). The system is anonymous and could be used on-line or with printouts during a timed session.

'HOW TO DO IT'

Prepare an assignment brief, which ideally contains guidelines to set the criteria to which the work is going to be marked. It is best suited for text-based