The Potential for Online Peer Assessment: Reflections from Two Modules

Fang Lou, Steve Bennett, and Trevor Barke

***Abstract*—The work described in this presentation relates to the evolution of peer assessment in first year undergraduate students in Life Sciences and Computer Science at the University of Hertfordshire. The objective of this research was to foster the development of higher order thinking skills and to deliver fast, effective feedback to learners in first year undergraduate modules in two different schools. In the first part of the project, the peer assessment of a laboratory report was introduced to first year Sports Science and Bioscience students. Bioscience students have been doing peer assessment for the last five years and their attitude has been very positive. However the Sports cohort appeared to be suspicious or hostile to the initiative. The methods used to improve the attitude of Sports' learners to peer assessment are reported here. The results showed that this year both cohorts have done very well and overall 88.2% of Bioscience students and 74.9% of Sports students answered online questionnaire agreed or strongly agreed to all the questions. In the second part of the project, we were able to show that Computer Science students following an E Media course benefited significantly by the introduction of peer assessment of previous cohort's work. In the first year an improvement of 6% was achieved in their final assessment and the following year a further 4%. This could only have been due to the peer assessment initiative. In the final section we discuss the potential for using peer assessment in online modules.**

***Index Terms*–Electronic voting system, evaluation and reflection, peer assessment, professionalism.**

I. Introduction

E-assessment is an issue of current interest to a wide community at present in Higher Education. A survey of E-Assessment in the UK, providing an overview and vision of the future landscape is provided by Whitelock & Brasher [1]. Peer assessment, where students mark each other's work, has also received a lot of recent attention in this context. A large amount of research has been undertaken on the beneficial effects of peer assessment on student motivation and ability to self-assess. Notably there have been three comprehensive meta-studies in recent times [2] – [4]. In previous research the use of online data gatherer was proved effective in encouraging reflection and feedback [5]. It was also found that peer assessment with Electronic Voting Systems (EVS) was useful, fair and beneficial to learners [6] – [7]. Often EVS is used in multiple choice tests or interactive quizzes as a means of enhancing learning and teaching [8] and to provide the opportunity for deeper learning [9].

In this project, first year undergraduate learners following course in schools of life sciences and computer science used peer assessment to help improve their reflection and higher order thinking skills. Here we describe how we were able to change the negative attitude of some learners to peer assessment.

Some of the findings were briefly reported as a conference presentation [10].

Manuscript received August 1, 2012; revised October 12, 2012. This work was supported by the Learning and Teaching Institute, University of Hertfordshire.

F. Lou is with the School of Life Sciences and the Learning and Teaching Institute, University of Hertfordshire, Hatfield, Herts AL10 9AB, UK (e-mail: F.Lou@herts.ac.uk)

S. Bennett and T. Barker are with the School of Computer Science, University of Hertfordshire Hatfield, Herts AL10 9AB, UK (e-mail: s.j.bennett@herts.ac.uk, t.1.barker@herts.ac.uk)

II. School of Life Sciences

A. Background Information

In the School of Life Science first year Bioscience students take a module Human Physiology, and Sports students take a module Foundations of Human Physiology. Both modules have laboratory practical classes and involve writing up a full laboratory report. Peer assessment of the full laboratory report had been used for four years in the Bioscience module Human Physiology. Students did very well and were very positive about this activity [5].

Last year peer assessment was introduced to the sports module Foundations of Human Physiology, but the attitude towards it of learners in the sports cohort had been shown to be worse than those following Bioscience modules to the initiative. Many students did not see the benefits of this activity and felt it was not their job to mark other students' work. They said they were uncomfortable of 'taking marks away from another student'. Due to the lack of motivation and incentive, many students sat in the marking session indifferently, and put a random mark without much thinking, and with no comments. This resulted in huge amount of moderation having to be done by staff, and the objectives of peer assessment were not achieved.

In order to overcome this poor attitude, a number of steps were taken this year to improve the Sports module, based on the discussions with the module delivery team and students who did the module last year. The changes are described in the next section. The Bioscience module had been kept as it was for the fifth year.

B. Steps Taken to Motivate Students in Peer Assessment

In 2011-12, we modified the peer assessment process in several aspects to motivate students.

First of all, we promoted peer assessment positively and right from the beginning. During the Induction week, each

module coordinator had a 15 minutes slot to introduce the module to the first year students. The peer assessment was mentioned as a highlight of the module (Foundations of Human Physiology) with a direct link to the students' future career – they would need to evaluate other people's performance in their professional job as a sport scientist or sport therapist.

Secondly, to reduce the potential hostility and fear towards peer assessment, we decreased the weight of the full lab report in the total module grade from 20% to 10%. The previous high stake assignment had become a medium stake assignment. So students felt it less threatening.

Thirdly, 5% of the total marks for the report were allocated to how well a student acted as a marker (1% for attending the marking session, 1% for putting down marks, 1% for reasonable marks, 1% for writing comments and 1% for good comments). This is to encourage students actively participating in the marking session.

Finally, 5% of the total report marks were allocated to the reflection and feedback by students via an online data gatherer. Many students may have written the report, submitted, and never seen the report again, apart from gaining a mark for it from the lecturer. We anticipated that the incentive would encourage students to read the feedback comments on their report and improve their next report.

C. Workshop to Further Motivate Students and Explain the Details

A one hour workshop was given to students after they finished the laboratory class and before they submitted the lab report. The workshop included:

1) Asking students' initial thoughts on peer assessment (using EVS - electronic voting system)
2) Rationale of peer assessment (including past data)
3) Details of the assignment
4) How to write a full laboratory report (including references)
5) General marking criteria for a laboratory report
6) Asking students' thoughts again (using EVS)

D. Submission and Preparation

Students were requested to submit an online copy of their report (for plagiarism detection and moderation), as well as a hard copy for peer assessing (with their student registration number only). Detailed specific marking criteria were carefully prepared.

E. The Marking Session

Two hours were allocated for the marking session. Students were given a report to mark along with a printed marking criteria and a marking sheet for Optical Marking Reader. A lecturer led the marking session with marking criteria on PowerPoint slides and another lecturer patrolling the room to answer questions (the number of staff needs to be increased with increased number of students, of course). An example was looked at using a visualizer. Students were encouraged to write comments for the marks they award/deducted according to the detailed marking criteria. After the marking session, the marked reports and the marking sheets were collected.

The marking sheets were scanned with the optical reader, and marks collated and announced with the students' registration number, and the marked reports were returned.

F. Reflection, Feedback and Evaluation

Students were given a username and a password for logging online to answer a series of open and closed questions on their thoughts on peer assessment after they received their marked report. The questions included what they had learned, and their opinion of the fairness of the marks they received. If they thought they were not marked fairly, they were asked to provide details of how differently they should be marked according to the marking criteria.Statistics Chi-square test and t-tests were applied.

III. Results

A. Importance of Motivating Students

As mentioned above, the first step we took was to promote peer assessment right from the Induction week with the Sports students. To find out how students were motivated towards peer assessment by the changes we made to the module, the answers to the electronic voting system (EVS) question at the beginning of the workshop were compared between academic years 2010-11 and 2011-12. At the beginning of the workshop, students were asked 'What did you think when you learned that we are going to do peer assessment?'. This question was answered anonymously using EVS handsets. Table I shows the percentage results of the students' response to EVS question:

TABLE I: Comparison of Students' Initial Thoughts on Peer Assessment between 2010-11 and 2011-12:

EVS answering choice:	2010-11	2011-12
1. Glad to have a go	0.0	13.0
2. Curious to find out how it works	0.0	35.2
3. Would prefer lecturers to mark it	80.0	48.1
4. Not comfortable with the responsibility	20.0	3.7

Data shown in the table are percentages. Number of student answered: 50 for 2010-11 and 54 for 2011-12, $p<0.001$ (Chi-square test).

The significant differences among the categories of the answers between the two academic years were mainly due to the promotion prior to the task. It can be seen that the initial barrier of fear and hostility was successful removed by the modification to the module during the Induction week.

B. Effectiveness of the Workshop

TABLE II: Comparison of Students' Thoughts before and After Workshop in 2011-12

EVS answering choice:	Before	After
1. Glad to have a go	13.0	27.5
2. Curious to find out how it works	35.2	7.8
3. Would prefer lecturers to mark it	48.1	51.0
4. Not comfortable with the responsibility	3.7	13.7

Data shown in percentage. Number of student answered: 54. $P<0.001$ (Chi-square test)

To test how well the workshop had prepared students for

the peer assessment, the same question was asked again to the students close to the end of the workshop (What do you think now about peer assessment?) and the results are shown in Table II:

The significant differences among the categories of the answers before and after the workshop were solely due to the workshop. It can be seen that the workshop was successful in clarifying the assignment, but did not change mind of those who preferred the lecturers to mark.

C. Marking Session

There was a huge improvement in the marking session for the Sports module. Most students took the peer assessment very seriously and put down reasonable marks and a full page comments. The lecturers moderated the marking before releasing the marks. There were very little changes made. The average score for marking was 4.7 ± 0.5% (out of 5%), which is 94%.

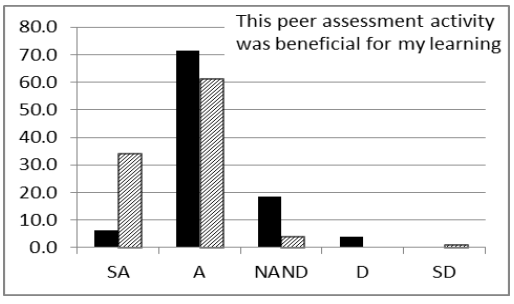
D. Quantitative Findings from the Online Data Gatherer

The online data gatherer gave us the possibility to collect students' opinion after they have received their marked reports. This year the rate of students submitting a full laboratory report, attending marking session and submitting online evaluation and reflection is listed in Table III below.

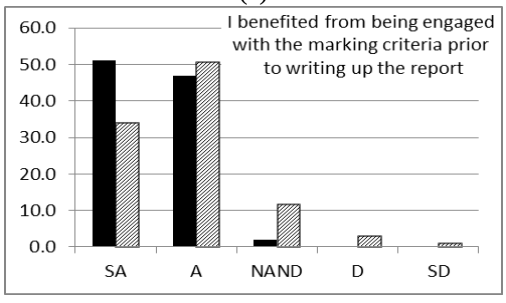
TABLE III: RETURNING RATE OF THE TWO MODULES

	FOHP (Sports)	HP (Bio)
Number of students	81	273
Reports submitted	79	255
Attended marking session	80	240
Online reflection	49	103

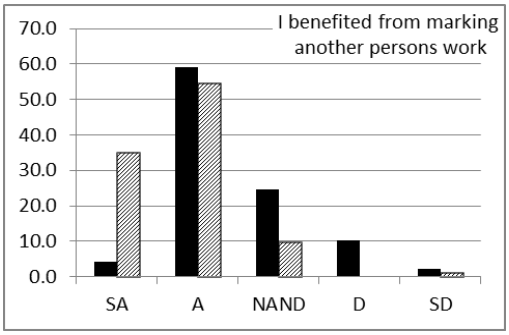
The online data gatherer has two parts. Part one consisted of closed question, meaning the answers were in numerical values (1 to 5), representing Strongly Agree (SA), Agree (A), Not Agree Not Disagree (NAND, or neutral), Disagree (D) and Strongly Disagree (SD). Figure 1 (a-f) show the answers from the closed questions. Part two consisted of open questions and collected students' written comments.



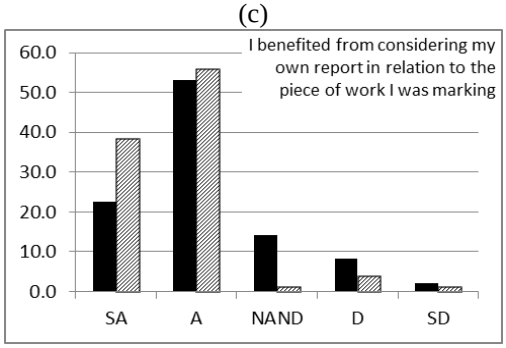
(a)



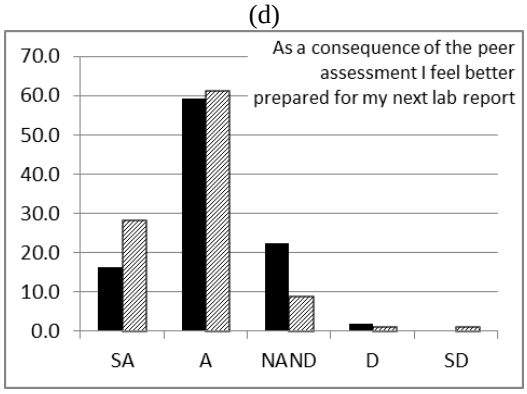
(b)



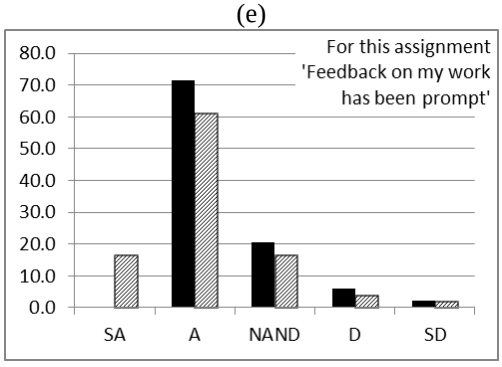
(c)



(d)



(e)



(f)

Fig. 1. Students responses to the questions indicated in each graph. Data were normalized to the total number of students responded (see Table III for details). SA: strongly agree; A: agree; NAND: not agree, not disagree (neutral); D: disagree; SD: strongly disagree. Light shaded columns: Bioscience students in Human physiology module; Dark shaded columns: Sports students in Foundations of Human Physiology module.

It can be been seen that majority of the students agreed or strongly agreed with all the questions. The only question that sport students agreed more than Bioscience students was 'I benefited from being engaged with the marking criteria prior to writing up the report'.

Averagely, 88.2% Bioscience students and 74.9% Sports students answered agree or strongly agree to all the questions. There has been a significant change in the attitude from the last year's peer assessment.

IV. School of Computer Science

A. Background Information

The E-Media Design module for first year Computer Science students is related to the development of a range of skills in the area of electronic media design. Many students following this module were not performing well on their practical work despite significant effort on the part of the tutors.

It was our hypothesis that the reason for the poor performance on the practical test related to the inability of some learners to internalize the assessment criteria required for the creation of a well-designed website. In short they were in possession of the necessary individual skills to perform the task but were unable to put them together in order to produce a coherent website with suitable form, function, content and aesthetics as required in the learning outcomes. Based on our findings from the MSc course reported previously [6] we decided to use a combination of Electronic Voting System (EVS) and a modified form of peer assessment in order to approach this problem.

B. The Assignments

Assignment 1 was a multiple choice test intended to test the theoretical component of the module. Assignment 2 was an assessment of the practical work of the previous cohort. Marks were awarded based on the closeness of a learner's mark to that of the tutors. This was intended to reinforce in learners the need to assimilate the tutors' criteria for the quality of the web sites. In order to do well, learners were required to understand details of the marking criteria, analyze, synthesis and apply the criteria in new contexts and to evaluate design decisions. Assignment 3 related to the design of a website according brief. Students were to prepare part of the website in advance and to complete the website later under examination conditions.

In the first year of the project (2010/2011) we were able to show that the initiative led to an increase of 6% in the mean performance of learners [7]. Despite this improvement, an amount of negative feedback from learners was received.

Some changes in our procedures were necessary in the following year (2011/2012), in order to address the students' issues. Firstly we changed the weighting of the feed-forward exercise. Reducing its value from 25% to 10%. In this way, it retained its importance, owing to its credit bearing nature, but the effects on the student's overall grade was much reduced. Secondly there was a huge rewriting of the scoring rubric.

V. Results (Quantitative Findings)

TABLE IV: PERFORMANCE BY LEARNERS ON ASSIGNMENT 1 AND 3 OVER A THREE YEAR PERIOD.

Test	Mean% 2009 Cohort	Mean% 2010 Cohort	Mean% 2011 Cohort
Assignment 1	59% (n=290) (SD 16.6)	56% (n=215) (SD 17.5)	54% (n=231) (SD 17.9)
Assignment 3	58% (n=277) (SD 18.4)	64% (n=218) (SD 14.3)	68% (n=221) (SD 16.3)

An ANOVA showed $p>0.05$ for the scores 2009; $p<0.001$ for 2010 and $p<0.001$ for 2011.

Data was collected and tabulated assignment 1 and 3 in each year. The data was summarized and taken into SPSS for statistical analyses. Table IV below presents a summary of the mean performance of learners on assignment 1 and assignment 3 over the three year period.

Assignment 1 was a multiple choice test. The items for the assignment in the three years studied were selected from the same item bank. The questions for each test were moderated by an expert who confirmed that the difficulty level of the assignments were more or less the same. The results from assignment 1 then provided a convenient base line or control for the level of the students over the three year period. After our first intervention in 2010, the mean score for assignment 3 was significantly higher than for assignment 1 (56% and 64%) ($p<0.001$). After our revised presentation in 2011, a further increase of 4% from the previous year and 10% from 2009 results was observed.

VI. Discussion

The work described in this paper was concerned with two related projects on the topic of peer assessment. The first project showed the adaptation of peer assessment in the marking of laboratory reports in sports science. The second project was able to show the benefits of the peer assessment of previous cohort's work using EVS clickers.

Findings from the Life Science project were able to show that the attitude of learners resistant to the peer assessment approach could be significantly improved by a careful procedure of selling the idea to learners and linking it to professionalism. The online data gatherer provided students a strong incentive to read the comments written by the assessor and reflect on their skills of writing a report. This is a huge improvement than a report being written, handed in, got a mark, and forgotten about. The ability of evaluation and reflection is extremely important in academic learning [11]. The data gatherer also helped to collect the students' opinion and enabled the tutors to follow up any issues concerned.

The findings from the Computer Science project were encouraging. Our initial problem of poor performance by learners on the practical test has to all intents and purposes been solved. The significant increase in performance over the three year period coupled with the fairly consistent base line strongly supports a genuine improvement in performance. Anecdotal evidence also supports this result. The quality of work submitted in assignment 3 strongly reflects the increase in marks awarded. It is of much higher quality as attested by the moderator who suggested we needed to add additional challenge future assignments as the average mark was getting too high. Interestingly the area of most improvement tended to be in the students achieving average marks. This may of course be due to a ceiling effect, and intend to investigate the possible reasons for this in future work.

In the 2010 presentation of the module, where, despite a significant increase in performance, we received a large amount of quite negative feedback from students. In 2011 we revised our approach in consideration of this. We made a far better attempt at selling the idea of the peer evaluation idea, linking it to professionalism and evaluation. We provided a richer and more informative rubric and increased the use of the EVS system. The result was that there was far less

controversy with the revised approach. There was clear evidence from the questionnaire data that students seemed to be buying in. In addition, there was absolutely no negative feedback related to the approach.

The benefits of peer assessment seem to relate to developing higher order thinking skills and the ability to internalize and reflect upon what is necessary in an assignment. In future work we intend to use peer assessment with our online learners. Obviously the use of EVS as in the first project and face to face as in the second would not be possible. We are currently investigating the use of Adobe Connect [12] in order to achieve this. In future work we will present examples of students' work to learners using Connect and require them to evaluate and discuss online. It is hoped that in this way the benefits of peer assessment will be available to these learners.

ACKNOWLEDGMENT

The authors thank the Learning and Teaching Institute, University of Hertfordshire, Hatfield, Herts AL10 9AB, UK for the financial support to the project. F. Lou thanks the teaching teams, especially Chris Benham and Mike Roberts for their support during the peer assessment.

REFERENCES

[1] D. Whitelock, and A. Brasher. "Developing a Roadmap for e-assessment: which way now?" in *Proceedings of the 10th CAA Conference 2006*, Loughborough University, 4/5 July, 2006.

[2] K. Topping, "Peer Assessment Between Students in Colleges and Universities," *Review of Educational Research*, 1998, vol. 68, no. 3, pp. 249-276.

[3] N Falchikov and J Goldfinch, "Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks," *Review of Educational Research*, 2000, vol. 70, no. 3 pp. 287 – 323.

[4] M. Zundert, D. Sluijsmans, and D. V. Merriënboer, *Unravelling Peer Assessment, and Instruction*, vol. 20, issue 4, August 2010, pp. 270-279 .

[5] B. Helen, F. Lou, and M. Russell *Peer assessment – Educationally effective and resource efficient. Blended Learning in Practice*, 2011, May, pp. 21-35.

[6] T. Barker and S. Bennett, "Marking Complex Assignments Using Peer Assessment with an Electronic Voting System and an Automated Feedback Tool," *International Journal of E-Assessment*, 2011, vol. 1, no. 1, pp. 2045-9432.

[7] S. Bennett and T. Barker, "The Use of Electronic Voting to Encourage the Development of Higher Order Thinking Skills in Learners," in *Proceedings of the International Conference on Computers and Assessment CAA 2011*, Southampton, July, 2011.

[8] M. B. Russell, *Using an electronic voting system to enhance learning and teaching, Engineering Education*, vol. 3, no. 2. 2008.

[9] S. W. Draper, "Catalytic assessment: understanding how MCQs and EVS can foster deep learning," Special Issue, Editor Whitelock, D., *British Journal of Educational Technology*, vol. 40, no. 2, pp. 285 - 293.

[10] F. Lou, S. Bennett, and T. Barker, *Using peer assessment and technology to support learning. Abstract. Learning and Teaching Conference*, University of Hertfordshire, 2012.

[11] D. R. Sadler, "Evaluation and the improvement of academic learning," *Journal of Higher Education* , 1983, vol. 54, pp. 60–79.

[12] Collaborate Solutions. [Online]. Available: http://www.collab8.com.