

**Keele University Innovation Project Report**  
**Web Support for Student Peer Review**  
**Stephen Bostock,**  
**Computer Science,**  
**July 2002**

|  |   |
|--|---|
| Introduction: student peer review.....           | 2 |
| Potential benefits to students as authors.....   | 2 |
| Potential benefits to students as reviewers..... | 2 |
| Potential benefits to tutors.....                | 2 |
| Previous experience: 1999/2000 .....             | 3 |
| Innovation project 2000/01-2001/02.....          | 4 |
| Objectives.....                                  | 4 |
| Constraints.....                                 | 4 |
| Version 1, 2000/01 .....                         | 4 |
| 2000/01.....                                     | 5 |
| Results.....                                     | 5 |
| Student evaluations .....                        | 5 |
| Version 2 .....                                  | 6 |
| 2001/02.....                                     | 6 |
| software improvements .....                      | 6 |
| 18 student evaluations .....                     | 6 |
| Best and worst aspects, student views.....       | 7 |
| Conclusions and recommendations.....             | 7 |
| Version 3 .....                                  | 7 |
| Appendices: Version 2 screens .....              | 8 |

## Introduction: student peer review

The assessment of student work is normally performed by tutors. However assessment can also be performed by student themselves and by their peers. This project concerns the assessment (or review) of student work by other students, either

- Formative of a draft or prototype, or
- Summative for a contribution to grading (with tutor moderation).

The assessment consists of one or both

- Quantitative (a % mark)
- Qualitative (text of constructive criticism).

### Potential benefits to students as authors

Students in large classes with hard-pressed tutors do not receive as much feedback on their work and performance as would help them learn. Student review provides extra feedback a tutor cannot provide. It is less expert feedback but probably as or more useful than can be provided by computer quizzes, which cannot process written work or software products of students.

Secondly, in the work described the student authors received multiple reviews (4 or 5). Although each may be less expert than the tutor, between them they may provide a wider perspective on the work.

Thirdly, in the work described here, the reviews use explicit and fairly detailed criteria, identical to or based on the final assessment criteria and the learning outcomes. Constructive criticisms should therefore help improve final performance.

### Potential benefits to students as reviewers

When students are assessors it can give some engagement with the assessment criteria. They may better understand the assessment criteria by having used them. If these criteria are also discussed or negotiated a further sense of ownership can be engendered. Being an assessor encourages students to perform the self-assessment needed to manage own learning and improve their own work for assessment. It also encourages a responsibility for their peers' learning rather than the default of competition.

Evaluation is a high order thinking skills and where (as here) it is one of the module learning outcomes, peer evaluation practises the skill to be assessed as an outcome summatively.

Finally, at the fundamental level of values, HE can be seen as an induction into a scholarly community where anonymous peer review is a key process for quality control.

### Potential benefits to tutors

Many of us would like to be able to provide more feedback to students - research is clear that it will improve their performance – but with larger classes and other demands on staff time it is often not possible. It might also be added that some of us find marking

large amounts of student work an uncongenial activity. What better than having other students mark the work, if it could provide useful feedback?

One effect that quickly becomes clear as a scheme is developed is that much greater clarity and detail in the assessment criteria is needed – enough so that even weak students can use them effectively. In the long run this is a benefit to the tutor as well as the student. Discussion of the criteria with the student body can become a productive (perhaps the most productive) activity for improving performance. Students will understand assessment criteria better and perform better.

The potential drawback of peer assessment for tutors that quickly became apparent was the burden of administration.

## Previous experience: 1999/2000

The core practical work (and 50% of the module assessment) in the *Multimedia and the Internet* module on the MSc in IT in Computer Science is the development of a web site. Every student has their own web space, visible within the Keele domain. Drafts or ‘prototypes’ of student web sites were submitted part way through the module and final web sites (accompanied by a report) were submitted at the end. Students were instructed to have a home page at a specific address. In 1999/00 both prototype and final web sites were reviewed by student peers. With 38 students I assigned them into author-reviewer so that each student reviewed 5 prototype web sites, I then sent emails inviting the authors to review the sites and giving the web addresses of the sites and of the web form to use in submitting the review. A standard web form using the cgiemail script was used. This sent the review data (text and numbers) to myself and to the author. I repeated the process for the final submissions.

### Results

The administration of the process was time-consuming and error-prone. Both I and some students made mistakes which had then to be corrected. For example, a review might be sent to the wrong author. 35 of 38 students did the formative reviews of text and percentage grades as requested. But when it came to the summative reviews of final submissions only 22 performed them – this was in the exam revision period and some student ‘went home’ to revise. An additional problem, foreseen in the design, was that the authors were not anonymous to the reviewers because they typed their email address into the form so that it would be sent to the appropriate author.

### Student evaluations

Of the 16 students who responded to an evaluation form, most said that the (reviewer) anonymity had allowed their criticisms to be ‘ruthless’ (honest), and hence more valuable. Also, as reviewers, seeing other students’ work had been valuable. The text criticisms of the prototypes were more valued than the percentage marks given.

On the other hand, the timing was difficult. They said they needed longer between receiving reviews and the final submission so as to use the criticisms. Some students complained of the time reviewing had taken.

Many anxieties were expressed about peer summative grading. The plan was that the average of the final student peer grades would be used, with tutor moderation of any awkward cases. Given the level of student concern about the need for moderation, all final student work was remarked by myself. Of course, there was no gain in staff time for the summative assessment.

This exercise confirmed many advantages but also that anonymity was essential, preferably for authors as well as reviewers, and that multiple reviewers were needed to provide balance and avoid difficulties of weak student reviewers. This meant that the administration performed manually was essential but needed to be automated. Automation would remove tutor errors and might also prevent student errors.

## Innovation project 2000/01-2001/02

### Objectives

The aim was to develop software to administer anonymous peer review, notifying students by email and collecting reviews by web forms. It should allow monitoring of the reviewing process by the tutor and archiving of the reviews. The intention was that the software would be generic, supporting any type of student review of paper or web materials, with any number of students as long as they had email addresses to receive instructions and access to the web to complete the review submission web form.

The tutor is to provide

1. Student authors' and reviewers' emails (in this case the same set, but in principle different sets of students)
2. The items to be reviewed (possibly as URLs of web sites or documents, otherwise by a code number)
3. The number of reviews to be performed per author
4. The criteria to be used by reviewers – these will appear on the form to be completed
5. The type of feedback required: text and/or a grade

Once the data is entered, the tutor initiates the review: the emails being sent and the form being made available. It was not intended that the software would automatically prevent late reviews after a deadline.

### Constraints

Anonymity is needed for reviewers and authors. The system must be secure from fraud or student error, only the correct reviews must be sent to authors, and only one review accepted for each reviewer-author pair. There should be equal reviewing loads for reviewers. The allocation of reviewers should avoid pairs of students who review each other's work, to mitigate against collusion. Finally, the software must run on a Keele web server, and therefore use Perl scripts.

## Version 1, 2000/01

2000/01

There were 68 students in the module. As part of the preparation for peer assessment they did a practice review on previous student work and then compared that to my marking of the work. This year I arranged for 4 formative and 4 (moderated) summative reviews per author, rather than 5 in the previous year. Reviews submitted were identified by a random code number *plus* the reviewer's username, which were entered on the web form and checked by the software against the list of legitimate reviewer-author pairs. The reviews themselves were emailed to authors and also turned into web pages for tutor to check.

### Results

Most formative and summative reviews were performed as requested. Some student errors in completing the review form meant that some students received too few, or wrong, reviews. The system was clearly not yet foolproof enough.

As in the previous years, the summative reviews were, in the end, not used and the work was completely re-marked, partly to see the accuracy of student marking as part of the project. The summative marks are summarised below for both years.

1999/00

Mean of students marks: 64%

Mean of tutor marks: 63%

Correlation with = 0.45

The (up to 5) marks per author had a mean range of 11% and a mean standard deviation of 7%.

2000/01

Mean of students marks: 63%

Mean of tutor marks: 63%

Correlation with = 0.59

The (up to 4) marks per author had a mean range of 13% and a mean standard deviation of 6%.

My conclusion is that average student marks are accurate but individual markers are very variable. It would be feasible to use student marks where the variation on a particular piece of work is low, and moderate them where it was high.

### Student evaluations

34 students returned the evaluation form.

Was the practice marking exercise useful? 88% said Yes

Was discussion of the assessment criteria useful? 87% said Yes

Were the reviews you received done professionally? 57% Yes, 21% No

On the reviews of prototypes:

58% said they were useful or very useful for making improvements

Were students happy with moderated summative peer assessment? 61% said Yes, but often cautiously, with caveats.

Should we do the peer assessment exercise next year? 79% said Yes

Clearly most students valued the exercise and their main concern was its use for summative marking.

## Version 2

2001/02

In 2001/02 there were 60 students on the module at Keele plus 55 pursuing the course with another tutor in Sri Lanka. Whereas in previous years students had been given a choice of the topic (content) of their web site, this year the topic was specified but different for each student: a tourist information web site for a specific country. Only formative reviews of the prototype stage were required – given the anxiety felt by students in previous years I decided that even moderated summative peer assessment was not, on balance, worth the effort. There were four reviews per author and this year the quality of the reviews themselves were assessed by me, so as to encourage all students to take reviewing seriously and give credit for the reviewing effort. This was consistent with a learning outcome of evaluation skills. Of the 4 reviews, 2 were marked at random to give 10% of the module mark (leaving 40% for the web site and 50% for the exam).

### Software improvements

Improvements to the software were

- Batch input of student lists where previously students were entered one at a time
- The security was improved: a unique code was built into the URL of the review form sent to the reviewer, and then used to by the software to identify the review, check it was not yet submitted before emailing it to the author and storing it as a web page for the tutor. Thus the identification of the author and reviewer no longer depended on the reviewer's entries in the form, all they had to do was use the URL as it was emailed to them. As a result, no reviews were mis-filed and very few were not completed (except for 2 absentees).

Some example screens from the software are appended:

[A List of author-reviewer pairs](#)

[The Criteria for coursework assessment](#)

[A Form for review submission](#)

[A List of Keele reviews submitted](#)

[An example review](#)

### Student evaluations

Of the 18 evaluation forms returned from the Keele cohort, only 6 found *all* the reviews they received were done ‘professionally’; this was worse than in the previous year! Most (13) found the reviews received were useful.

They were split evenly on their views about possible use of reviews for summative use (which did not happen this year). 69% recommended using peer review in a following year. They had similar likes and dislikes as in the previous year:

Best aspects:

- Getting constructive comments
- Learning to design by evaluating others’ designs
- Seeing other students’ work
- Clarifying the assessment criteria

Worst aspects:

- The time taken doing reviews
- Reviewing was not sufficiently anonymous (they felt)
- Receiving some poor reviews
- Prototypes they reviewed were too incomplete
- Being ‘kind’ to others when writing review comments

### Conclusions and recommendations

Formative student reviews of work in progress (prototypes) have been found to be valuable for students both as authors and as reviewers. Students receive extra feedback, constructive criticism, on their work. The reviewing process is a valuable learning activity. Both aspects clarify the assessment criteria.

There is a benefit in having multiple reviewers (4 or 5) and anonymity for reviewers and authors if possible. It may be worth ‘policing’ the quality of reviews by tutor assessment of them for a portion of the grade, the evidence is unclear in this case.

To administer multiple anonymous reviews with moderate or large numbers of students needs software support. The software used in 2001/02 is being completed as version 3 to support peer reviews in any discipline, of paper or web materials.

## Version 3

The final version ideally will be

1. Capable of managing multiple review events in separate courses
2. Moved to a server allowing wider access
3. Have a web interface and login for tutors to
4. Upload student lists, and edit them
5. Generate the web form from criteria, for formative and/or summative purpose
6. Generate and check author-reviewer pairs
7. Generate emails to reviewers
8. Allow a tutor to monitor the reviews as they are submitted by web forms.

## Appendices: Version 2 screens

[List of author-reviewer pairs](#)  
[Criteria for coursework assessment](#)  
[Form for review submission](#)  
[List of Keele reviews submitted](#)  
[An example review](#)