An Effective System for Electronic Peer Review

MICHAEL DE RAADT

Department of Mathematics and Computing and

Centre of Research on Transformational Pedagogies

The University of Southern Queensland

Toowoomba, Queensland, Australia

deraadt@usq.edu.au

MARK TOLEMAN

Department of Information Systems and

Centre of Research on Transformational Pedagogies

The University of Southern Queensland

Toowoomba, Queensland, Australia

markt@usq.edu.au

RICHARD WATSON

Department of Mathematics and Computing

The University of Southern Queensland

Toowoomba, Queensland, Australia

watsonr@usq.edu.au

Electronic peer review can empower instructors of large courses to produce rapid feedback, promote social interaction and encourage higher order learning by students. But what are the payoffs for educators? Do students recognise the benefits of such a system? Foundation Computing is one of the largest courses at the University of Southern Queensland. A system of electronic submission and peer reviewing with instructor moderation is now being used in this course. This system is innovative and unique and delivers benefits to students, instructors and the University. This system has been evaluated, proven successful and is being considered for wider use.

Key words: peer review, rapid feedback, social interaction, higher order learning

Introduction

Peer review, online or in the classroom, is not a new pedagogical practice. Much peer review literature relates to assessing peers' contribution to work completed in a group. The system presented in this paper is different since it creates peer review relationships for each assessment item. This allows students to give reflective feedback immediately after submission of their own work and reduces the delay from submission to feedback receipt. In online peer-review research the focus is often on online discussion with involvement in discussion as a means of assessment (Prins *et al*, 2005). Kurhila *et al* (2003) present a system which allows students to create HTML documents which are peer-reviewed but not peer-assessed, whereas student reviews in produced by the system presented in this paper are used, with instructor moderation, for assessment purposes. The Calibrated Peer Review (CRP) system (Chapman, 2006) is a tool which encourages higher order thinking and offers reduced marking load through peer review; these same benefits are sought from the system presented here. CPR is limited to essays where system reported here allows students to submit and review documents of differing formats related to the various applications that are studied during the course. CPR requires an initial training of students in an attempt to calibrate the peer reviews they produce. The peer reviews created under the CPR system are subjective and not guaranteed to be reliable or reproducible. The system presented here requires no student training and uses objective criteria and instructor moderation to ensure validity of marks gained through peer review.

Aims and Research Questions

Foundation Computing is a computing concepts course that covers applications skills (one third) as well as theoretical aspects (two thirds). The course is run by domain experts and services students from various disciplines. Each year there are a total of approximately 1000 enrolments in the course. Two thirds of students enrolled are external (studying at distance). All students have access to online facilities.

In the previous incarnation of the course, two paper assignments were assessed. The paper based process could take six weeks from submission to feedback receipt. Management and marking created a great cost burden for the University. Paper assignments were handled in two areas of the University before being logged to markers and returned by the reverse route. It is estimated that introducing electronic submission can be seen as saving the equivalent of at least one person for one day per week when considering up to 600 students can be enrolled in a single semester.

The course was altered with the intention of delivering improved outcomes for students. Such aims lead to research questions examined in this paper.

- **Regular assignments and rapid feedback**

  Using electronic means, submission and feedback delivery is almost instantaneous. An electronic peer reviewing system with moderation allows students to receive feedback, from multiple sources, in greatly reduced time. It is therefore possible to offer eight regular assignments consisting of focussed, learner-centred tasks. This raises research questions: Do electronic submissions and rapid feedback benefit students with limited study time? (RQ1) Do they provide flexibility for highly motivated students by allowing them to work ahead and still receive feedback within a reasonable time? (RQ1) Do students see these benefits? (RQ3)

- **Authentic and personal tasks**

  Focussed tasks can force students to discover how the real world of computing affects them personally. Such tasks include designing a resume or hypnotically building a computer for their needs. Do students appreciate such tasks? (RQ4)

- **Higher Order Thinking**

  Higher order thinking can be encouraged through synthesising finished works and evaluating the works of others (Bloom, 1956). If students are required to complete reviews of other students' submissions do they achieve higher order thinking? (RQ5)

- **Social interaction**

  Students are able to communicate through the course website. A list of users is displayed which is rarely empty of students who are online. The focus of course communication is the course Bulletin Board, where questions are asked and

answered. Students electronically review other students' work and give feedback using a purpose-built system. Do these aspects establish an online community of learning, reducing student isolation and further encouraging higher-order thinking (Brook and Oliver, 2003)? (RQ6)

- **Shift workload from marking to teaching**

  The electronic submission and review systems, which replaces the previous paper based assignment approach, greatly reduces cost of managing assignments, both in time and people-power. This benefits teaching staff, freeing them to focus on other tasks such as consultation. But does the system allow marking staff to be applied in other forms of teaching? (RQ7)

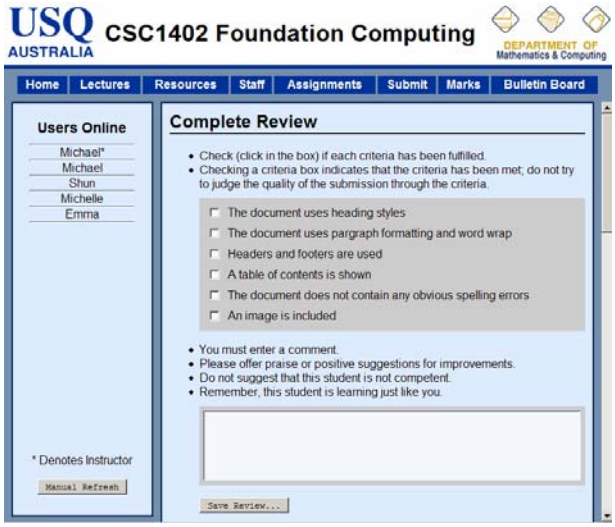


Figure 1: Review Facility

Description of the Innovation

Students submit eight assignments over the semester, seven of which are peer reviewed. During the semester students create word processing documents, spreadsheet documents, presentations and HTML documents. Students are introduced

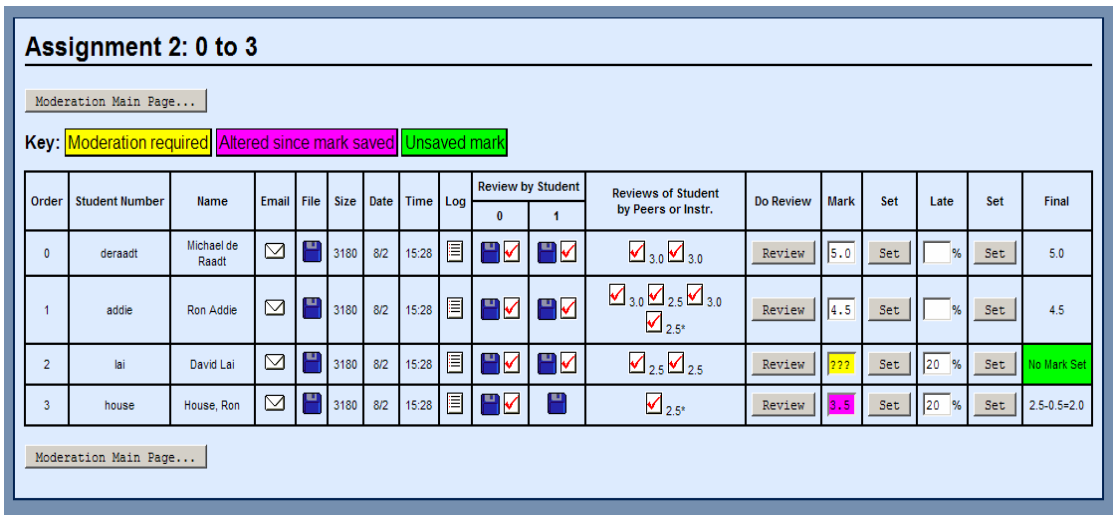


Figure 2: Moderation Facility (false names used)

to an application then use that application to explore theoretical parts of the course. Students submit assignments electronically through a Web facility. To prevent incorrect submission, students must submit assignments in order. As students will see other students' solutions when completing reviews, only one submission is permitted per assignment. Students can (repeatedly) practice using the submission and review systems with a pseudo-assignment to build confidence before actual submission and reviewing. As deadlines are regular and frequent, heavy late penalties apply and late assignments are only accepted for five days after the due date. Students can work ahead by submitting early. Some students have completed all assignments during the first half of the course.

After submission, students are automatically allocated two other students' works to review. The allocation algorithm is sufficiently complicated so students cannot anticipate who will review their work. Students are awarded marks for completing reviews. The first four students to submit must wait for other assignments to be submitted before they can review. The system then notifies these students by email when reviewing is possible.

Each student downloads and anonymously reviews the work of two other students based on set criteria using the facility shown in Figure 1. Students are required to give a comment and are encouraged to give praise and constructive criticism. Student reviews are used as the basis for marking. The success of the system relies on students being able to make consistent and fair reviews. To achieve this, criteria must be objective and easily discerned. Students must be able to correctly recognise whether criteria have been met in the work of a peer, even if they have failed to meet the criteria in their own work. Criteria focus on completeness of task rather than judgement of quality; this reduces ambiguity and increases consistency among reviewers. Students are asked to make encouraging observations and suggest improvements in comments that accompany the review. Students are made aware of the criteria in the assignment specification, so they know how they will be reviewed before they submit.

Where two reviews differ according to the criteria, an instructor moderates the assignment to give a definitive mark. Approximately one in every ten non-conflicting

reviews are also moderated. If a student is unhappy about the review they received, they can flag it for instructor moderation. Instructors utilize the same form students use for completing reviews, but employ a tabular interface to launch the moderation process as shown in Figure 2. Instructors performing moderation have access to a large pool of information about each submission, the student who submitted it and the students who reviewed it. This aids in detecting plagiarism and checking for possible collusion. Over the semester about half of all submissions are moderated. The number of instructor moderations is tracked so that students who have not been moderated previously can be targeted.

Outcomes and Evaluation

The system was piloted during the summer semester (November to February) 2004-2005 with 160 externally enrolled students. One of the most pleasing results was the speed with which students received their first feedback (usually a peer review). Half of all submissions received feedback in 1hr 21min or less. There were exceptions as some students submitted well in advance of the deadline and at least three submissions must accumulate before any reviewing can commence.

Table 1:Time from submission to first feedback

Min	3min
1st Quartile	34min
Median	1hr 21min
3rd Quartile	4hrs 19min
90%	14hr 11min

In the last weeks of the semester, students were asked to participate in a voluntary survey. The survey was conducted online. 90 students responded to the survey which is equivalent to 62% of the students who were submitting the last two assignments at around the same time. As well as collecting information about age, gender and an experience self-assessment, the survey consisted of 24 statements to which the students were asked to indicate their agreement on a five point Likert scale. Some of the statements were positively phrased such as "I would be happy to use the same submission and review facilities in other courses". Others were negative such as "I did not receive enough support to complete the assignments". Some statements were designed to estimate participant willingness to adopt such technologies and such an approach to assessment in future (see Venkatesh *et al*, 2003). Other statements were used to elicit the students' understandings of the potential pedagogical benefits of

using this system. Participants had the opportunity to add comments which are examined in de Raadt, Toleman and Watson (2005).

Demographics

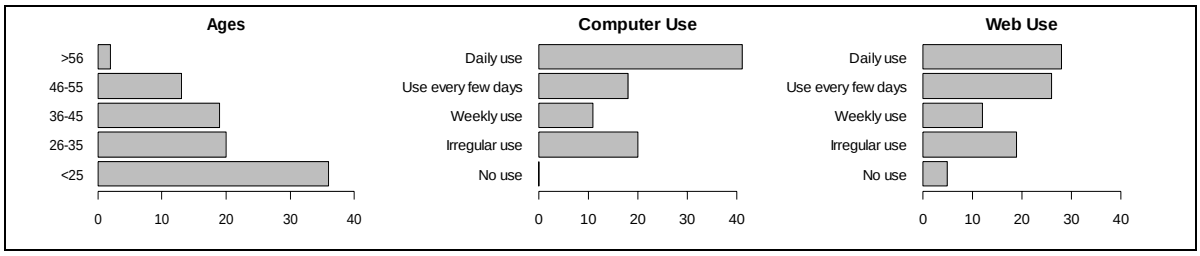
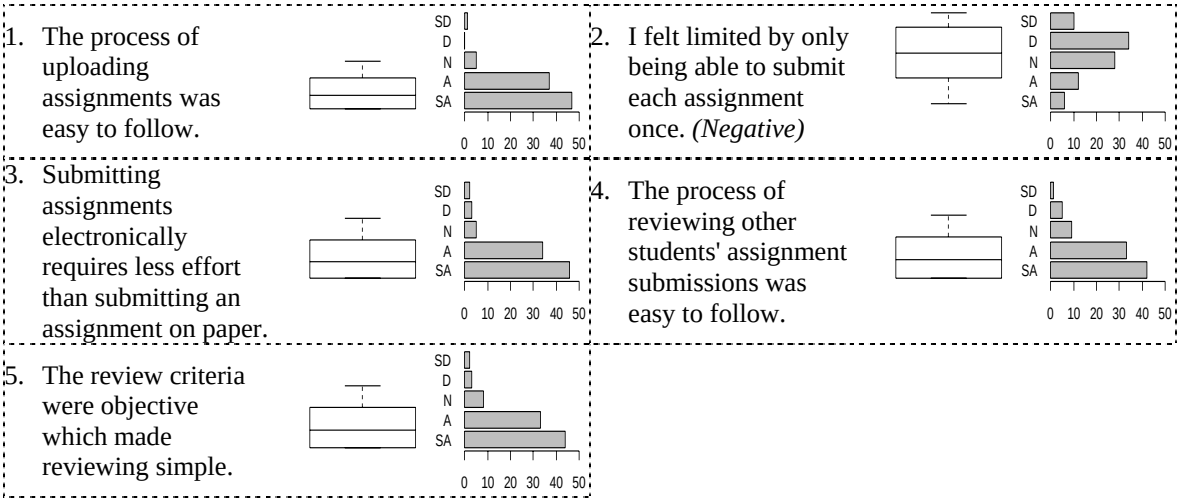


Figure 3: Age and Computer Experience

Three quarters of the participants were female and one quarter was male, which was representative of the students enrolled in the course. Figure 3 shows number of student responses to questions of age range and level of computer and Web use prior to commencement of the course. Students who participated in the survey covered a wide age range; 40% of participants were 25 or under, the remaining participants could be considered as 'mature aged'. All participants had used a computer prior to starting the course. Very few participants had not used the World Wide Web before studying in the course.

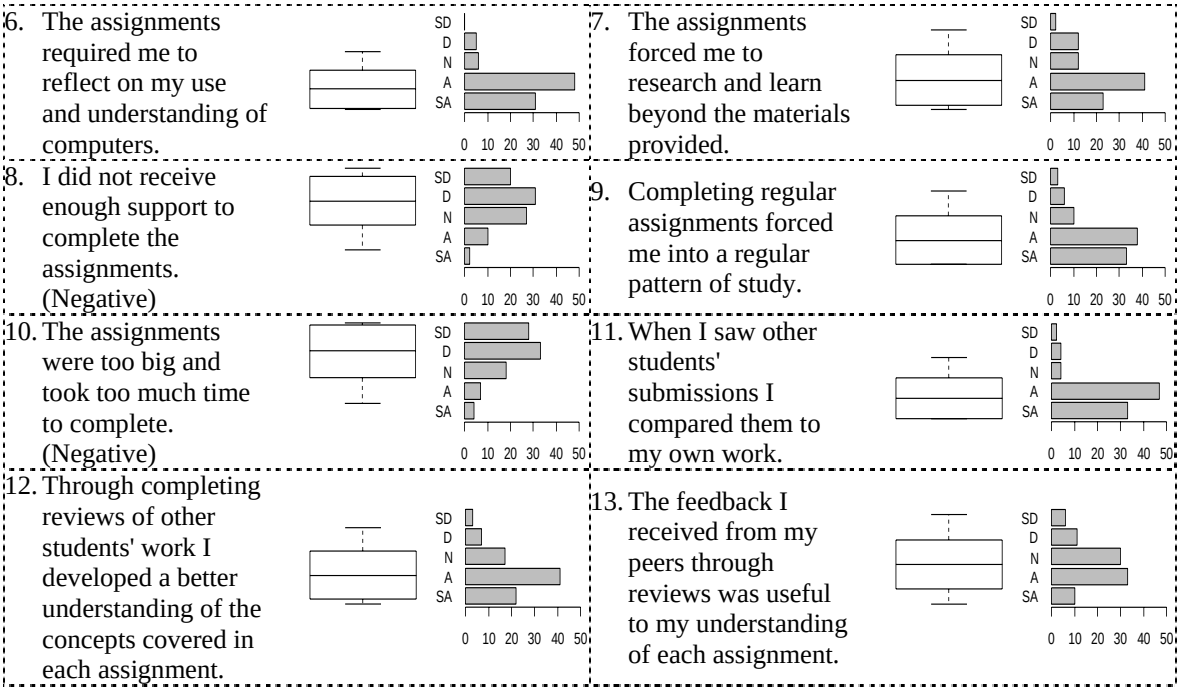
Each of the 24 survey statements is shown below with a summary of responses. Possible levels of agreement were strongly agree (SA), agree (A), neutral (N), disagree (D), or strongly disagree (SD). In summaries the number of responses to each level of agreement is revealed in the bar chart on the right. On the left of this the distribution of responses is exhibited using a box and whisker plot with a mean response indicated by the horizontal middle line of the box, a single standard deviation shown by the box around this and two standard deviations shown by whisker lines.

Ease of Use



Participants indicated the process of submitting and the process of reviewing and applying review criteria (in statements 1, 4 and 5 respectively with agreement SA+A over 83% in each) were easy to follow. This indicates that the facility was well designed and implemented. The vast majority of participants recognised that there was less effort in submitting electronic assignments compared to a paper submission (statement 3, SA+A=89%). A neutral response was given about being limited by one submission per assignment (negative statement 2, D+SD=49%, N=31%, SA+A=21%). There was a mechanism for dealing with accidentally submitting incorrect documents, which required instructor intervention, but students were not told about this explicitly.

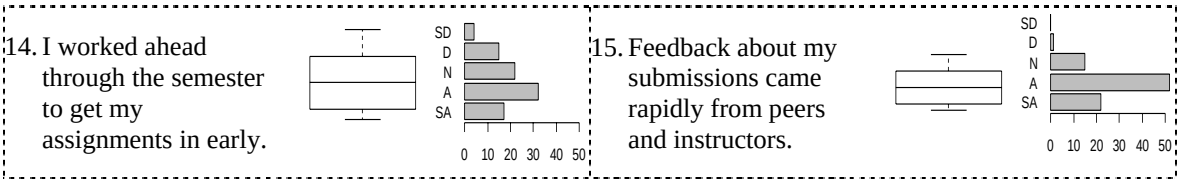
Benefits to Learning



Without prompting most participants saw the system's learning benefits. The assignments caused the greater majority to reflect on their computer use (statement 6, SA+A=88%). Completing regular assignments forced students into a regular pattern of study; most participants recognised this (statement 9, SA+A=79%, N=11%, D+SD=10%). A strong correlation was found between students with little prior computer/web use and those who reported being forced into a regular pattern of study (Pearson $r=0.39$, $p<0.05$). This may be because they were forced to use a computer more regularly to complete assignments, submit and review. The same group was also correlated to a strong response in reflecting on their understanding of computers (Pearson $r=0.28$, $p<0.05$). Participants indicated completing reviews caused them to compare their own work with their peers' (statement 11, SA+A=89%) and through completing reviews, most students developed better understanding of concepts (statement 12, SA+A=70%, N=19%, D+SD=11%). The assignments were successful in forcing most participants to learn beyond the materials (statement 7, SA+A=71%, N=13%, D+SD=16%). Participants were not prompted to suggest if they felt this was a positive or negative aspect of the course. Many participants were neutral about whether there was enough support provided to complete assignments (negative statement 8, D+SD=57%, N=30%, SA+A=13%). The assignments were designed to encourage students to construct their own learning in several areas. Instructor

assistance was provided through several means, but this response may indicate that many students expected a more prescriptive teaching approach. Participants indicated the assignments were not too large (negative statement 10, D+SD=68%, N=20%, SA+A=12%), however there were a reasonable proportion who did feel the assignments were too large. It is difficult to determine if these participants were responding from a learning perspective or from a workload perspective. One aspect which became apparent was that many participants were neutral about, or gave little value to, the reviews of their work conducted by peers (statement 13, SA+A=48%, N=33%, D+SD=19%). This may be because students felt the assessment of their work by peers carried little expertise (as indicated by some student comments) or that the amount of feedback or aspects reviewed were not sufficient using the review system.

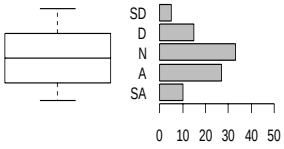
Time



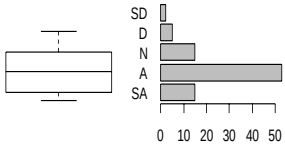
Some students did work ahead; most submitted on the due date. In comments from students, some said they liked the idea of being able to work ahead, but did not take advantage of it. Of course, working ahead was not required, but from students' comments, some participants misconstrued this statement as suggesting they should have been working ahead and were expressing that this was not required. The vast majority recognised the response time benefits (statement 15, SA+A=82%, N=17%, D=1%).

Quality

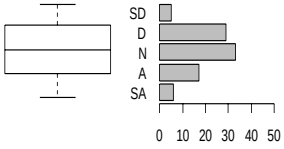
16. Feedback on my submission was as good or better than what I would expect on paper based assignments marked by hand by an instructor.



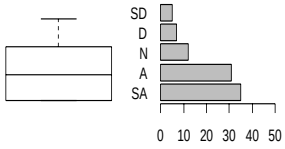
17. The reviews I received were fair and consistent.



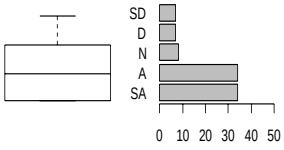
18. I would rather receive marks from instructors only. (Negative)



19. Help was available if I had trouble using the system.



20. I would be happy to use the same submission and review facilities in other courses.



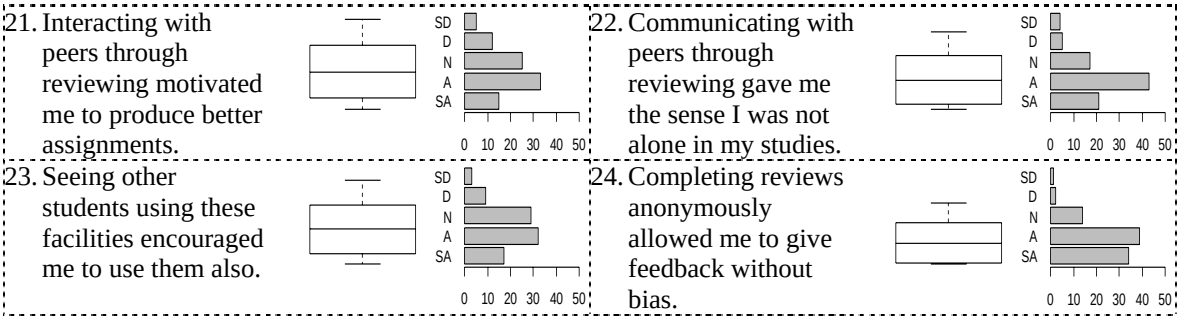
Participants expressed varied opinions on whether feedback they received was superior to assignments hand-marked by an instructor (statement 16, SA+A=41%, N=37%, D+SD=22%). While the system delivers rapid feedback from more sources and other learning benefits through reflection, it does not necessarily deliver better feedback than a hand-marked paper assignment. One goal of the system is to match the feedback quality possible with hand-marking. Part of this feedback quality relies on students providing accurate reviews. Participant's responses varied widely about peer feedback compared to instructor-only feedback (negative statement 18, D+SD=38%, N=37%, SA+A=26%). It was noted during moderation that some students did not give useful comments or made errors during reviews. It is possible for a student to complete a review without viewing the submitted document by blindly checking all criteria boxes and leaving a comment such as, "Well done." Instances of this were discovered due to conflicting student reviews. One student had reviewing marks removed where it was obvious they had not made an effort to review their peer's work.

Most participants felt the reviews they had received were fair and consistent indicating the set criteria were objective and easily discerned (statement 17, SA+A=76%, N=17%, D+SD=8%). Most students felt there was help available if needed in using the system (statement 19, SA+A=73%, N=13%, D+SD=13%). Very

few students requested assistance during the initial running of the system. This approving response is perhaps also indicative of the system's perceived ease of use.

Perhaps the biggest endorsement of this approach is most participants agreeing that they would like to use the same facilities in other courses (statement 20, SA+A=76%, N=9%, D+SD=16%). Some participants commented that this should definitely happen.

Interaction



One aspect of general technology adoption (commonly applied to business and industry settings) involves potential adopters witnessing colleagues, especially superiors, using the proposed technology. A majority of participating students in this study felt encouraged by seeing their peers use the system, but there was a large group who were ambivalent (as suggested by statement 23, SA+A=54%, N=32%, D+SD=13%). This may be because all participants were studying externally, not having face-to-face contact with their peers. This may also be an indicator of an approach to study that differs from an industrial/business setting or it may indicate that participants did not see their student peers as superiors. One might think that a student, knowing their peers would be reviewing their assignment, might be motivated to produce a better assignment submission. Again a majority indicated they were motivated by this, however there was a large group who were neutral or not motivated (statement 21, SA+A=53%, N=28%, D+SD=19%). Perhaps participants were indicating that they would have produced assignments of high quality if only an instructor was marking them.

In order to encourage fair and unbiased reviews it is important that anonymity be maintained between the reviewer and the reviewee. The reviewer must be confident

to deny marks or provide criticism where appropriate. The system did not identify reviewees to reviewers or vice-versa, however, in several cases, even though submission of anonymous work was encouraged, students identified themselves within the documents they submitted. Despite this, most students acknowledged they felt free to give comments without bias (statement 24, SA+A=81%, N=16%, D+SD=3%).

Participant responses showed it is possible to create a community in the setting described here (statement 22, SA+A=71%, N=19%, D+SD=10%). A feeling of community is essential to avoid effects of isolation that a student can feel when studying independently at a distance.

Conclusions

Electronic peer-reviewing can deliver benefits to students, instructors and the University.

Students benefit by:

- receiving rapid feedback from multiple sources (RQ1),
- being free to work ahead and still receiving timely feedback (RQ2),
- practicing skills relevant to them personally (RQ4),
- evaluating other students' work and reflecting on their own work thus achieving higher order thinking (RQ5),
- learning how to share documents,
- gaining experience in using online computerised facilities, and
- perhaps most importantly, becoming more involved in the course, feeling less isolated and potentially further encouraging higher order thinking (RQ6).

Most students recognise the benefits of electronic peer reviewing (RQ3).

Instructors can experience reduced marking load; such time was used to increase the availability of instructors for contact and the authors believe this encouraged better learning outcomes (RQ7).

The University benefits through reduced manual management of assignments at various levels. This could potentially reduce costs for the University.

Feedback from participants showed students found the system easy to use. Electronic submission requires less effort than submitting paper assignments. Regular assignments can force students into a regular pattern of study, particularly those with little previous experience in the content area. By providing clear and objective criteria, fair and consistent peer-reviews can be achieved. Peer reviewing can contribute to a community focussed environment and help to lessen isolation of students studying at distance. Most students would happily adopt this approach in other courses. Using the system does place constraints on students. For example, only a single submission is possible. Students need to feel comfortable about the submission and review processes. Students need to be assured that when problems arise they can be resolved by an instructor easily and without penalty. The reviews and comments made by peers are not valued highly by many students. Encouraging greater accuracy may raise this value. Students must feel they are being assessed regularly by course instructors, even if it is not apparent for all assignments. Students are not likely to be motivated by what their peers think when determining their willingness to use such a system.

The benefits of the facilities and assessment approach discussed here are not apparent to all students. It is incumbent on instructors to explain the benefits of the peer-review process to students. Examiners should admit that reducing marking load is a motivation for using peer-review and assessment systems. It is then possible to claim that time saved can be applied to more productive teaching. The benefits to the student from reviewing such as evaluation, rapid feedback and other aspects of the approach could be explained to students so they are more aware of why such systems are used.

Following the initial success of this approach, it has been adopted as the continuing model of assessment for Foundation Computing with great enthusiasm by the team that manage this course. The approach is being applied to both external and on-campus students. With greater confidence in the system, and to reward the effort put in by students who participate in it, the value of all eight assignments has been

increased from 40% to 56%. The peer-review system shown here is being considered for use in other courses, including an introductory programming course. The next phase of system development is to modify the system so it can be applied to new courses rapidly.

References

Bloom, B.S. (Ed) (1956) *Taxonomy of Educational Objectives Handbook 1: Cognitive Domain*. New York: Longman, Green & Co.

Brook, C. & Oliver, R. (2003) Online learning communities: Investigating a design framework. *Australian Journal of Educational Technology*, 19(2), 139-160.

Chapman, O.L (2006) *The White Paper: A Description of CPR*. Available online at http://cpr.molsci.ucla.edu/cpr/resources/documents/misc/CPR_White_Paper.pdf (Accessed 23rd Feb, 2006).

de Raadt, M, Toleman, M & Watson, R (2005). Electronic Peer Review: A large cohort teaching themselves? *Proceedings of the 22nd Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE2005)*. Brisbane, 4-7 December, 2005.

Kurhila, J., Miettinen, M., Nokelainen, P., Floreen, P. & Tirri, H. (2003) Peer-to-Peer Learning with Open-Ended Writable Web, *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education*, 173-178, ACM Press.

Prins, F.J., Sluijsmans, D.M.A., Kirschner, P.A. & Strijbos, J. (2005) Formative peer assessment in a CSCL environment: a case study, *Assessment & Evaluation in Higher Education*, 30(4), 417-444.

Venkatesh, V., Morris, M., Davis, G. & Davis F. (2003). User Acceptance of Information Technology: Toward a Unified View, *MIS Quarterly*, 27(3), 425-478.