Effective Peer Assessment for Learning Computer Programming

Jirarat Sitthiworachart
Department of Computer Science
University of Warwick
Coventry CV4 7AL, UK
Tel +44 24 7652 2368

jirarat@dcs.warwick.ac.uk

Mike Joy
Department of Computer Science
University of Warwick
Coventry CV4 7AL, UK
Tel +44 24 7652 3368

M.S.Joy@warwick.ac.uk

ABSTRACT

Peer assessment is a technique that has been successfully employed in a variety of academic disciplines, and which is considered to be effective in developing student's higher cognitive skills. In this paper, we consider the results of applying novel web-based technology to the delivery of peer assessment in the context of an undergraduate computer programming course, and discuss the benefits of this approach.

Categories and Subject Descriptors

K.3.1 [**Computers and Education**]: Computer Uses in Education – *Collaborative learning (peer assessment)*.

General Terms

Management, Measurement, Human Factors, Languages.

Keywords

Peer assessment, peer review, programming languages, UNIX.

1. INTRODUCTION

Assessment measures student ability when used summatively, and helps students to learn when used formatively. Topping et al.[5] define peer assessment as "an arrangement for peers to consider the level, value, worth, quality or successfulness of the products or outcomes of learning of others of similar status". Peer assessment is a technique which is generally considered to be effective in promoting students' higher cognitive skills [4], since students use their knowledge and skills to interpret, analyse and evaluate others' work in order to clarify and correct it [1,2].

Experiments in deploying peer assessment have been performed in a number of different academic disciplines, such as peer assessment of academic writing in an educational psychology module [5], designing creative lessons on student teachers [18], group presentation of physical therapy student [13], and in

problem based learning (PBL) of electrical and information engineering student [3]. All of them reported both benefits and problems, which occurred during the process. This paper describes our experiences with a novel web-based peer assessment system deployed on a large class of undergraduate computer science students studying computer programming. We discuss the benefits of our approach, and consider unresolved issues in relation to experiences reported using peer assessment in other disciplines.
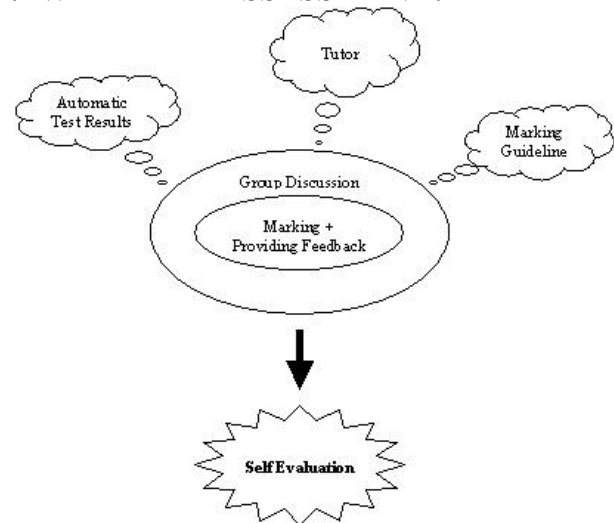
2. WHY PEER ASSESSMENT?

Figure 1 Peer assessment activity

The use of peer assessment is claimed to enhance students' evaluative capacities [4], which improve the quality of their subsequent work [5,6]. The software we developed contains components, which incorporate the use of

- automatic test results (so that students gain an appreciation of the correctness of the programs they are viewing);

- strict marking guidelines; and

- tutor support throughout the process.

There are three important activities in this peer assessment process, each of which provides benefits to students in improving their learning, as follows (Figure 1).

- *Group discussion*: students have a chance to exchange their knowledge, express their own ideas, understand more about the assignment, and improve their interpersonal skills [7].

- *Marking*: when marking, students review and compare their own work with their peers' work. They analyse and evaluate other work, which leads to the development of evaluation and self-evaluation skills [4].

- *Providing feedback*: students correct other work and explain their arguments [1,2], which encourage the students to reflect. Furthermore, it can be of great value to the educator both for teaching and for assessing course outcomes [17].

Many academics have performed paper-based peer assessment, where students fill in marking and feedback sheets, and rotate the assignment to their peers [2, 3, 8]. Thus it is difficult to make peer assessment anonymous and to monitor students' marking. Also there is no award for the marking task [11], which leads to inconsistencies in the quality of marking.

In order to address these issues, we developed a modified peer assessment process. We implemented a web-based peer assessment system [9, 10], which we evaluated on 215 first year undergraduate students (189 male and 26 female). The chosen module was a first year UNIX programming module, which aims to give students a basic understanding of the UNIX operating system, and competence in programming using a UNIX shell. Students learn how to design and develop programs in the shell, which is a programming language that allows programs to be written in many styles. The purposes of performing the experiment in peer assessment were:

- to investigate the extent that peer assessment *in a programming course* promotes students' higher cognitive skills;

- to assess the accuracy of students' judgements during a peer assessment exercise; and

- to provide evidence that peer assessment in computer programming has a positive pedagogical effect.

3. WEB-BASED PEER ASSESSMENT
The success of an online exercise such as our web-based peer assessment depends on four steps: design, software development, deployment and monitoring.

3.1 Design
3.1.1 Process
In many peer assessment exercises, students only mark and provide feedback on other work. Some students do not perform this task seriously because it does not affect their mark. Therefore we should add another stage – mark quality of feedback (Figure 2) in order to encourage students to take the assessment seriously and make this marking process more effective. The average mark for this part of the exercise was 79%, indicating that students provided quality marking and feedback. This stage also helps students to develop their critical judgement skills, they can see how other student mark and provide feedback to compare with their judgement. Thus the more marking students did, the better

their own results became [11], and this is supported by students' comments:

> "Marking and writing suggestions, helps me to understand how it is supposed to be."

> "From marking other scripts, helps me criticise my own work. I am able to see different programs written by the people in the same course. It makes me think what the general idea is and something that I don't know yet. So I can revise or chase what I am missing."
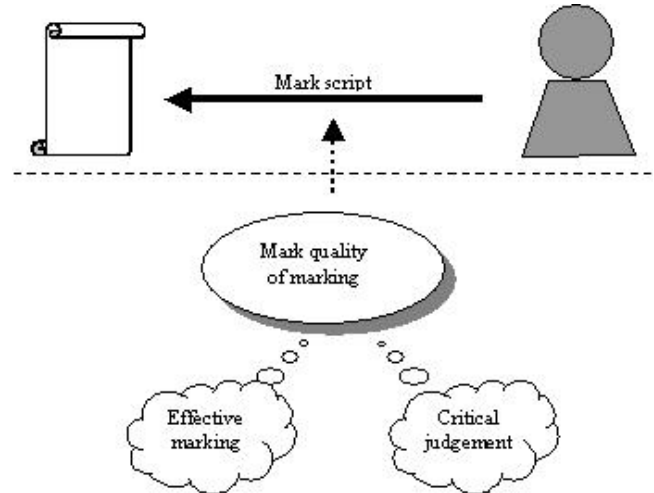

Figure 2 Marking process

3.1.2 Anonymity
Peer assessment should be anonymous, and this is supported by interviews with our students, in which 94% of students interviewed agreed, in order to avoid friendship marking and lessen embarrassment of marking friends' work. This aspect also has been reported by Segers et al.[2] and Davies [12], that anonymous peer assessment reduces the opportunity of collusion and biased marking. Our students remarked:

> "Peer assessment - anonymous is a good idea, because if it's not anonymous, I incline to give bad marks to people I know."

> "I don't want to do peer assessment exercise, if it is not anonymous."

3.1.3 Group discussion
Group discussion is an important factor in the peer assessment process. Student can share knowledge and learning [7], resulting in greater understanding of the assignment and the development of transferable interpersonal skills. In our research, we divided students into small groups (three students per group), each group consisting of students with a range of abilities. Each student was assigned three other students' assignments to mark from other groups (each consists of a mixed range of quality of assignments). They can discuss their marking with the other students in their group, who marked the same assignments. In the experiment, students were allocated timetabled laboratory sessions where they were able to talk together in a group, with tutors present to offer assistance if required. Students were positive about group discussion, as illustrated in the following responses:

"Group discussion helps me to understand more about the assignment as we had to go through each script in detail to make sure we could understand it."

"I enjoy working in a group. I like collaborating with others to decide what marks to give, which helped me learn more than one way of solving the programming problem."

With face to face discussion may arise the problem of students being unwilling to talk to each other, and feeling intimidated about asking questions. Furthermore the absence of some members of the group may complicate the process. Thus an anonymous communication system is recommended to solve these problems, where students can discuss online or leave offline messages to their anonymous group.

3.1.4 Mark scheme
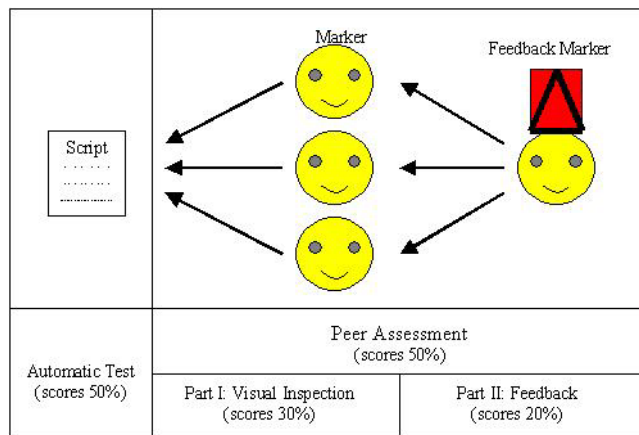
Figure 3 Mark scheme

When marking a programming assignment, we consider the correctness and quality of each program. We have an automatic test system, which runs each student's assignment against different test inputs to check the *correctness* of its functionality. Student peers then mark the *quality* of each program, a measure which is not strictly defined and which offers a variety of interpretations. Furthermore, the Unix shell allows for different approaches to solving a given programming problem, and there is no "template" against which the quality of the program can be marked against. Therefore in this peer assessment process, we weight the mark: 50% of the marks are awarded by the teacher

(automatic tests) and the remaining 50% are awarded by the students (Figure 3). Peer awarded marks are 30% for quality of program and 20% for quality of marking. All peer marks are based on three markers, the average of the three marks being calculated and used. The results of a post assignment questionnaire revealed that 65% of students were satisfied with their mark from the peer assessment, and 51% of students considered that the peer feedback they received was relevant and useful. Since students did not have marking experience, some of them hesitated to comment on other students' work. Thus some students gave only positive comments on the quality of the programs presented, the choice of utilities, the adherence to the assignment specification, and the comments in the programs, without offering solutions. This reason goes some way to explaining why 49% of students did not consider that the peer feedback they received was relevant and useful.

3.1.5 Marking criteria and guidelines

We initially chose to simplify as much as possible the marking criteria and the range of choices available to the students, supported by information from our pilot study. There are only three choices for each marking category, i.e. 'No', 'Partial', and 'Yes'. However, we found that 39% of students felt uncomfortable when assigning marks because of the small number of choices:

"The scale was not a very good one. How do I mark a script that has got most things right, but has a few defects? Do I mark it correct or partially correct? It seems harsh to take marks off people when they are quite close to being fully correct, but maybe not close enough to award them all the marks."

Therefore marking criteria should be clear and marking scale should offer appropriate choices for marking (a 5-point Likert scale is recommended – see Table 1) in order to help student in making an accurate and fair judgment [13]. Using specific marking criteria also helps student to understand what is expected of a good program, as the following response from student.

"When going through the marking criteria, it's always reminds me "what did I do". Seeing someone got a comment there, you didn't think it's appropriate or may should I should make my comment more detailed, make more sense for people who read the program."

Table 1 Marking criteria in Unix programming module

1. Comments are	unhelpful	○○○○○	helpful
2. The code are indented	inconsistently	○○○○○	consistently
3. Variable/function names are	inappropriate	○○○○○	appropriate
4. The code handles errors	inappropriate	○○○○○	appropriate
5. The program finishes with an appropriate exit status	never	○○○○○	always
6. The utilities have been selected	inappropriate	○○○○○	appropriate
7. The program is structured	poorly	○○○○○	well
8. Overall, following what the program is doing is	hard	○○○○○	Easy

Orsmond et al. [14] reported that students seemed unenthusiastic about creating marking criteria themselves. Therefore in this peer assessment process, marking criteria were set by the teacher. As students have different levels of knowledge and ability in marking, marking guidelines or things to consider should be provided to point out potential aspects of a good answer. We found that 95% of students understood the marking criteria and guidelines, indicating that the clarity was appropriate. In this peer assessment process, we also provided the students' automatic test results as well. Student can see the automatic tests, the expected output and the actual output, which helps them in marking.
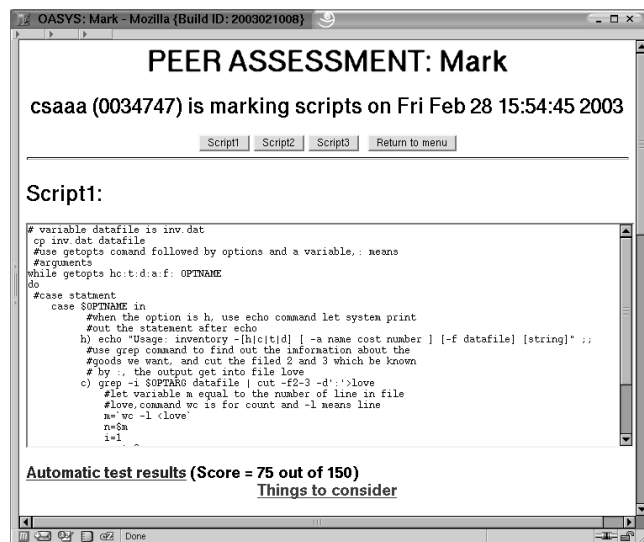
3.2 Software Development

Figure 4 Web-based peer assessment

In this peer assessment process, students submit assignments via the department's online submission system [15], and then evaluate each other's work through the web (Figure 4). Web-based peer assessment has advantages over ordinary peer assessment because students can be more critical due to the anonymity the system provides. Although each student initially marks in a laboratory environment (in order that support can be provided by tutors) they can finish the process wherever and whenever they choose before the marking deadline, thus allowing anonymity to be preserved. The processing of the marks is automated, and teachers can easily monitor the marking. Provision of online discussion and an offline messaging system helps student in discussing marking:

"I have experienced peer assessment before we worked in a group. After finishing marking one group then we had to swap this assignment to other groups. So it's not anonymous. This peer assessment is much better, we can mark online at home. So it's more convenient."

3.3 Deployment

One of the major factors in the successful implementation of peer assessment is to be careful in introducing it to staff and students, and to ensure they have a correct understanding of the objectives and benefits in enhancing learning rather than as a method of grading [4]. A few students misunderstand the purpose of peer assessment:

"Oh, I suppose when the system is set up perfectly, there isn't as much marking for lecturers to do, but that isn't a benefit for me, obviously."

Thus it is important to provide enough time to discuss the process with students at the beginning of the exercise. Moreover the following issues also should be considered.

- Taking the mystery out of the process, enabling students to appreciate why and how marks are awarded [16].
- Training students to rely on their own judgement and their peers, and to develop a belief that a tutor or lecturer is a coach, who supports and adjusts the decision that students make [18].
- During the marking process, tutors should be available for discussion about the process [17].
- Introduction of a tutor moderation system would be value to the peer assessment process to prevent cheating and enable the process to be seen to be fair [2].

3.4 Monitoring

It is inevitable that issues will arise during such an experiment, such as student absences, and students who – for whatever reason – do not fully engage with the process. Monitoring of the exercise is thus an essential part of the process, to ensure that no student receives an unfair mark or inappropriate feedback. Availability of appropriate basic statistical data, such as the standard deviation of marks for each group of markers, allows the teacher to intervene and – if appropriate – remark. In order to evaluate the accuracy of the marks awarded by the students, tutors also remarked all the students' pieces of work. There was a good correlation between the (expert) tutors' marks and those awarded by the students, though the latter were on average 17% higher.

4. DISCUSSION

Peer assessment is as much about learning as assessment, and student peers may not have adequate knowledge and experience to evaluate others' work, even when guidance and well-explained marking criteria are provided. The nature of the programming assignment (no model solution, no single "right" answer) was such that the opportunity was available for students to reflect on the technical material presented during the peer assessment process. Also there are some deeper issues from the students' perception, which it is difficult to see how they can be overcome without substantial and continual exposure to peer assessment processes (89% of students never have experienced peer assessment before). For example, more than 50% of students think that they are not qualified to be marker; only the tutor or lecturer is the expert marker.

"My script was marked by students, not somebody who is appropriately qualified and has been marking scripts for years."

"I don't like the idea of someone not qualified and possibly even knowing less about the subject than me, marking my work. I feel that the scripts should have been marked by lecturers."

However, many students have the different idea about this, such as the following response.

"Each marker has more time to mark individual pieces thoroughly, thus allowing him/her to give better feedback and

fairer marking than an automatic test-based system. It also gives marks which may be more aligned with the level of the candidate's knowledge and understanding – since the lecturer already knows his subject, whereas we are only seeing this for the first time!"

Some students did not believe their peers would mark fairly and consistency (although our results indicate this fear is misplaced).

"I don't know for sure exactly how well I did – I don't know how much I can trust someone who is at the same level as me."

"There is no consistency between markers, some are good and some are bad, they mark by their own standards."

Although marking is anonymous, it is difficult for student to assess their classmates. This also was reported by Sluijmas et al. [18] and Sivan [19].

"Sometimes I feel either being generous or harsh with marking. Very hard to do when you are marking someone who is technically in the same boat as you."

"At times I felt that if I mark a script too harshly, then I would ultimately affect their overall mark. And if I marked it too leniently, then they would be getting marks easily. So there was too much pressure on the marker."

5. CONCLUSION

We have discussed the use of a web-based tool, which has been evaluated with a large class of undergraduate students studying computer programming. The novel features of the tool – the awarding of marks for the students' marking itself, and facilities for anonymous marking – have demonstrably clear benefits. The use of very clear marking criteria, appropriate marking scales, and useful marking guidelines, are "good practice" which helps the process of peer assessment. With careful planning and implementation, students have a positive view of peer assessment and realize the educational benefits to them.

6. REFERENCES

[1] Segers, M., Dochy, F., "New Assessment Forms in Problem-based Learning: the value-added of the students' perspective", Studies in Higher Education, 26(3), 327-343, 2001

[2] Ballantyne, R., Hughes, K., and Mylonas, A., "Developing Procedures for Implementing Peer Assessment in Large Classes Using and Action Research Process", *Assessment & Evaluation in Higher Education*, **27**(5), 427-441, 2002

[3] McDermott, K., Nafalski, A., and Gol, O., "Active Learning in the University of South Australia", 30th ASEE/IEEE Frontiers in Education Conference, October 18-21, 2000, Kansas City, Missouri

[4] Fallows, S., and Chandramohan, B., "Multiple Approaches to Assessment: reflections on use of tutor, peer and self-assessment", *Teaching in Higher Education*, **6**(2), 229-246, 2001

[5] Topping, K., Smith, E., Swanson, I., and Elliot, A., "Formative Peer Assessment of Academic Writing Between Postgraduate Students", *Assessment & Evaluation in Higher Education*, **25**(2), 149-169, 2000

[6] Sluijsmans, D., Dochy, F., Moerkerke, G., "Creating a Learning Environment by Using Self-Peer-and Co-Assessment", *Learning Environments Research*, **1**, 293-319, 1999

[7] Andriessen, J., Working with Groupware: Understanding and Evaluating Collaboration Technology, Springer-Verlag London Limited, UK, 2003

[8] Purchase, H., "Peer Assessment: Encouraging Reflection on Interface Design", *23rd Australian Computer Science Conference: ACSC 2000*, 196-203, Australia

[9] Sitthiworachart, J., and Joy, M., "Web-based Peer Assessment in Learning Computer Programming", The 3rd IEEE International Conference on Advanced Learning Technologies: ICALT03, Athens, Greece, 9-11 July 2003

[10] Sitthiworachart, J., and Joy, M., "Deepening Computer Programming Skills by Using Web-based Peer Assessment", 4th Annual Conference of the LTSN Centre for Information and Computer Sciences, NUI Galway, Ireland, 26-28 August 2003

[11] Bhalerao, A. and Ward, A., "Towards electronically assisted peer assessment: a case study", *Association for Learning Technology journal (ALT-J)*, **9**(1), 26-37, 2001

[12] Davies, P., "Computerized peer assessment", *Innovations in Education and Training International*, **37**(4), 346-355, 2000

[13] Miller, P., "The Effect of Scoring Criteria Specificity on Peer and Self-assessment", *Assessment & Evaluation in Higher Education*, **28**(4), 383-394, 2003

[14] Orsmond, P., Merry, S., and Reiling, K., "The use of Exemplars and Formative Feedback when Using Student Derived Marking Criteria in Peer and Self-assessment", *Assessment & Evaluation in Higher Education*, **27**(4), 309-323, 2002

[15] Joy, M., Chan, P., and Luck, M., "Networked Submission and Assessment", *Proceedings of the 1st Annual Conference of the LTSN Centre for Information and Computer Sciences*, LTSN Centre for Information and Computer Sciences, 2000

[16] Brindley, C., Scoffield, S., "Peer Assessment in Undergraduate Programmes", *Teaching in Higher Education*, **3**(1), 79-89, 1998

[17] McGourty, J., Dominick, P., and Reilly, R., "Incorporating Student Peer Review and Feedback into the Assessment Process", 1998 FIE Conference, 14-18

[18] Sluijsmans, D., Brand-Gruwel. S. and Merrienboer, J., "Peer Assessment Training in Teacher Education: effects on performance and perceptions", *Assessment & Evaluation in Higher Education*, **27**(5), 443-454, 2002

[19] Sivan, A., "The Implementation of Peer Assessment: an action research approach", *Assessment in Education*, **7**(2), 193-213, 2000