



THE UNIVERSITY OF
MELBOURNE

Involving students in peer review

**Case studies and practical strategies for
university teaching**

Jon Pearce, Raoul Mulder and Chi Baik



This guide was developed for the University of Melbourne by Dr. Jon Pearce of the Department of Information Systems, Associate Professor Raoul Mulder of the School of Zoology and Ms Chi Baik of the Centre of the Study of Higher Education.

Permission is granted for copying, distribution and use by other institutions, with appropriate acknowledgement.

Available in electronic form from <http://www.cshe.unimelb.edu.au/>

© CSHE 2009
ISBN 9780734041432

Cartoon on page 5 by Nick Kim, with permission

Further queries regarding permissions and availability:

Centre for the Study of Higher Education
The University of Melbourne
<http://www.cshe.unimelb.edu.au>



CONTENTS

INTRODUCTION.....	2
1. WHAT IS STUDENT PEER REVIEW? A CONCISE REVIEW OF THE LITERATURE.....	3
BENEFITS OF PEER REVIEW	3
POTENTIAL IMPEDIMENTS TO IMPLEMENTING STUDENT PEER REVIEW	4
2. INCORPORATING STUDENT PEER REVIEW IN TEACHING: FOUR CASE STUDIES	7
CASE A: EXPERIMENTAL ANIMAL BEHAVIOUR	7
CASE B: PEER REVIEW IN MULTIMEDIA AND COMMUNICATIONS	8
CASE C: RESHAPING ENVIRONMENTS	10
CASE D: PROGRAMMING LANGUAGE IMPLEMENTATION	11
REFLECTION ON THE CASE STUDIES	12
3. REFLECTIONS ON IMPLEMENTING STUDENT PEER REVIEW	13
3.1 WHY USE STUDENT PEER REVIEW IN TEACHING?	13
3.2 STUDENT PERCEPTIONS OF THE PEER REVIEW EXPERIENCE	14
3.3 KEY CONSIDERATIONS IN IMPLEMENTING STUDENT PEER REVIEW	17
4. MANAGING STUDENT PEER REVIEW USING ONLINE TOOLS	23
CONTEXTS AND SOFTWARE CAPABILITIES	23
REFERENCES AND RESOURCES	28

Introduction

The concept of student peer review has gained increased attention in higher education in recent years, in line with the growing focus on peer and collaborative learning (Falchikov, 2005; Van den Berg et al.; 2006). Peer review is a form of peer assessment that has been used for over three decades in a wide range of disciplines from Architecture to Computer Science to Music, just to name a few.

The peer-review process involves collaborative learning in which students assess one another's work (usually against explicit criteria), and provide each other with feedback. The process is formative and has the aim of improving student learning while it is happening, in order to enhance understanding of the work involved and hence improve the quality of the final product (Somerville, 1993; Topping et al., 2000). This is distinguished from peer assessment for grading or summative purposes.

This guide has been developed as a resource for teaching staff and course/subject coordinators who are considering integrating student peer review in their subjects. While the case studies and tools described in this booklet are specific to the University of Melbourne, the discussion has broad applicability and relevance to other higher education institutions in Australia and abroad.

The guide is organised into four sections. The first section briefly reviews theoretical work on peer-review in higher education. This is followed by a section illustrating how academics in different discipline areas have taken diverse approaches to incorporating student peer review in their teaching, and a section discussing the benefits and issues in implementing student peer review. The last section presents an overview of online tools that are currently available to academics for managing student peer reviews. Together these sections aim to address three questions:

- *What* is student peer review?
- *Why* incorporate student peer review in subjects/courses?
- *How* can student peer review be implemented effectively in subjects/courses?

1. What is student peer review? A concise review of the literature

In a student-learning context, peer review is understood to mean the educational arrangement in which students consider or evaluate the value, quality or success of work produced by their fellow students and provide each other with feedback (Topping et al., 2000, p.150; Van den Berg, 2006, p341). Put simply, peer review involves students giving and receiving feedback on each other's work. It is also referred to in the literature as formative peer assessment, peer evaluation, peer response, or peer editing (Lundstrom and Baker, 2009).

The use of student peer review means a shift away from the traditional notion of assessment and feedback as being solely the role and responsibility of the teacher (Fallows and Chandramohan, 2001). Rather, peer review allows and encourages students to take an active role in managing their own learning. In this way, it is closely related to self-assessment as the skills required for peer review are similar to those required for self-review or self-assessment, which is critical for lifelong learning (Boud, 2000; Liu and Carless, 2006).

Peer review comes in various forms and has broad application. For over three decades, it has been used to evaluate a wide range of student work including written assignments, oral presentations, artwork and architectural designs, programming and code reviews, musical performances, as well as being used in various teamwork and capstone projects (Geringer, 2001; Topping, 1998)¹. It may involve a component that includes graded peer assessment of student work, graded assessment of the review, or a combination of both.

Students' level of involvement in the peer-review process may also vary from using the set criteria (or model answer) provided by staff in evaluating their peers' work, to negotiating and developing the criteria with staff before applying these standards to their peers' work (Brew, 1999). Through engaging in this process, students gain a better understanding of what the criteria are for good learning and this enhances their ability to select 'good evidence' (Biggs and Tang, 2007, p.187). The process thereby helps students to develop skills that facilitate their own learning (Dochy et al., 1999; Liu and Carless, 2006).

Benefits of peer review

Feedback is an essential part of the learning process. Effective feedback can motivate students, change their behaviour and improve their learning (Liu and Carless, 2006; Rubin, 2006; Biggs and Tang, 2007).

Through engaging in peer review and receiving feedback from a number of peers, students are exposed to a greater diversity of perspectives than just those of their tutor or lecturer. Indeed, the ability to ensure that students receive feedback from a number of different people is one of the main benefits of student peer review. As Rubin (2006), a Management lecturer notes, 'four independent student reviewers were able to provide much greater depth and breadth of feedback than I do alone' (p.391). Also, because students care about what their peers think about them, some authors (e.g. Falchikov, 2005; Gibbs, 1999) argue that students often pay more attention to peer feedback that has a social dimension than to a mark given by their tutor.

For teaching staff, implementing student peer review can be a way of ensuring that all students receive prompt and detailed feedback that can then be used for self-reflection and to make improvements to

¹ For a comprehensive typology of peer assessment in higher education, see Topping (1998)

their work (Liu and Carless, 2006). This is particularly beneficial in large classes where it may be difficult for the lecturer/tutor to provide detailed and timely feedback to all students. Rubin (2006) reports for example, that incorporating student peer review in his large Management course greatly increased his ability to provide students with substantive feedback.

In addition to the benefits from receiving feedback, numerous studies have highlighted the learning benefits derived from being involved in the process of reviewing and providing feedback on peers' work (see for example, Topping, 1998; Dochy et al., 1999; Falchikov, 2005; Sluijsmans, 2005; van den Berg et al., 2006; Lundstrom & Baker, 2009). Peer review of students' writing, for example, allows for meaningful interaction with peers, greater exposure to ideas, and new perspectives on the writing process (Hansen & Liu 2005, Lundstrom & Baker, 2009).

The review process requires students to analyse, review, clarify and sometimes correct each others' work. This can help to clarify and reinforce the reviewers' knowledge and understanding of the area and encourages the development of advanced critical thinking and higher-order cognitive skills (Topping, 1998; Gehringer et al., 2001).

The peer-review process can also enhance student learning by:

- building problem solving skills through identifying areas needing improvement and providing constructive suggestions (Dochy et al., 1999; Somerville, 1993);
- encouraging reflection and thereby promoting skills in self-assessment (Liu and Carless, 2006);
- enhancing greater meta-cognitive self-awareness (e.g. Topping, 1999; Liu and Carless, 2006);
- increasing student motivation by fostering a sense of responsibility and ownership for their peers' learning (e.g. Dochy et al., 1999; Warren and Cheng, 1997);
- promoting independent learning and reducing dependence on staff as 'the experts' (Brindley and Scoffield, 1998; Dochy et al., 1999);
- improving self-confidence (Brindley and Scoffield, 1998); and
- providing valuable experience and preparation for the professional workplace (Brindley and Scoffield, 1998; Biggs and Tang, 2007).

The literature on this topic also points to benefits to a range of social and communication skills such as verbal or written communication, negotiation skills, diplomacy and giving and accepting criticism (Topping et al., 2000). Engaging in peer review can enhance relationships in the group (Warren and Cheng, 1997) through increased interactivity self-confidence and empathy for others (Topping, 1999). It can also influence students' attitudes about learning from their peers – they may start to see other students and themselves as 'legitimate sources of knowledge' (Gehringer et al., 2005, p.321). This can in turn have a positive effect on the development of a collaborative and participatory learning environment (Fallows and Chandramohan, 2001).

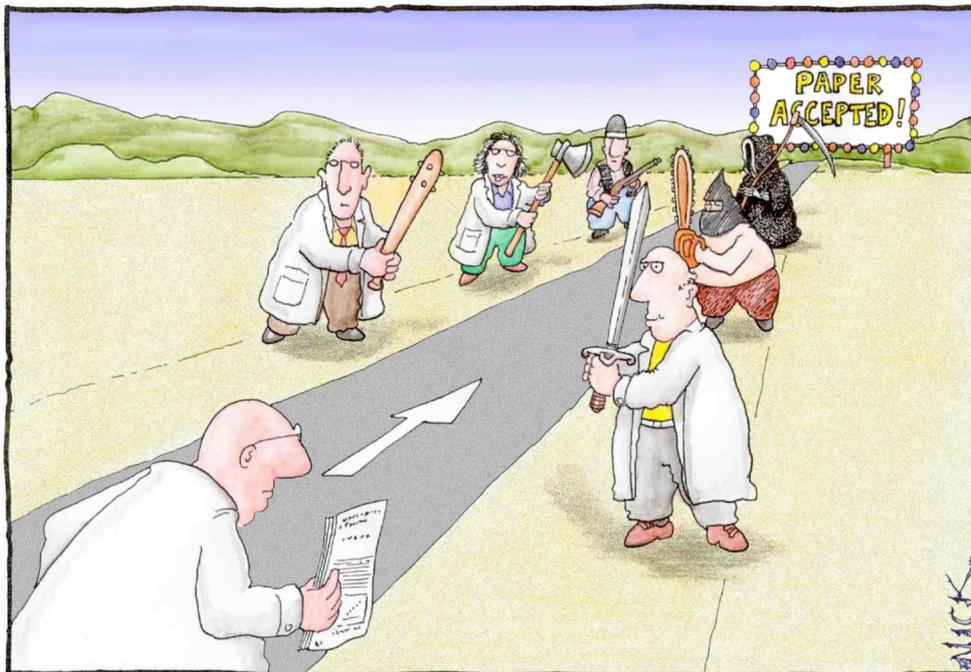
Potential impediments to implementing student peer review

Although numerous studies have highlighted the benefits of engaging in peer review, various authors have also identified several issues and potential impediments to implementing student peer review. Most of the difficulties discussed in the literature relate to peer grading of student work and to the assessment concerns of validity, reliability, bias and fairness. While these are generally not causes for concern in the formative types of peer assessment discussed in this guide, there are still a number of practical and potential impediments that need to be considered to ensure that students achieve the intended learning benefits of engaging in peer review.

An important consideration (and potential obstacle) in implementing student peer review is related to student perceptions and attitudes about the peer review process. Several studies have shown that

students may dislike evaluating another student's work. Biggs and Tang (2007) for example, note that some students resent being required to review and comment on other students' work, because they hold the belief that assessment is the 'teacher's responsibility'. This view is confirmed by Brindley and Scoffield's (1998) study in which the majority of students regarded assessment (and feedback) as solely the role of the tutor, and by Fallows and Chandramohan (2001) who also note that their students challenged the shift in responsibilities.

Another reason for students' discomfort with the idea of peer review is because they may lack confidence in their own ability to evaluate their peers' work. They may similarly doubt the competence of other student reviewers (Cheng & Warren, 1997; Sluijsmans et al., 2001). To deal with this concern, Fallows and Chandramohan (2001) advise providing guidelines or training for reviewers and discussing the rationale for, and benefits of, peer review. This will strengthen the students' awareness of the process and assist them to critique their peers' work. It may also add positively to students' overall satisfaction with the review process (Cheng and Warren, 1997; Dochy et al., 1999).



Most scientists regarded the new streamlined peer-review process as "quite an improvement."

Some students may be reluctant to engage actively with the peer review process because undertaking peer review of two or more students' work may be perceived to be overly time consuming, and they may feel that the 'cost' (in terms of time), outweighs the learning benefits they receive (Hanrahan & Issacs, 2001; Topping et al., 2000). Related to this is the question of whether students will take the peer-review process seriously if it does not count for marks. Hanrahan and Issacs (2000) point out two possible issues with this: first, reviewers may not put a great deal of effort into the process; and second, the recipients may not take the process or feedback seriously either. The authors thus suggest awarding marks for the quality of feedback provided by student reviewers. Evidence indicates that this can positively contribute to how engaged and committed students are to the task of reviewing (Hanrahan & Issacs, 2000; Geringer, 2001; Rubin, 2006).

Another important consideration for staff is whether they need to monitor the type of reviews given by students. Specifically, how will staff manage situations where students provide feedback that is overly critical and harsh, or where students provide little or no critique in their reviews (Rubin, 2006)? To minimize potential problems of this kind and ensure that all students benefit from the peer review process, Falchikov (2005) argues that staff need to model effective feedback and spend time training and teaching students how to review effectively.

From the perspective of teaching staff, a practical obstacle to implementing peer review concerns the time it takes to set up the process. Depending on how peer review is implemented, the organizational/administrative load associated with large classes can be significant, particularly if one of the aims is to protect the anonymity of reviewers (Hanrahan & Issacs, 2001; Brindley & Scoffield, 1998). Despite the time needed initially to set up the process, Rubin (2006) notes that time savings 'iteratively increase with each new round of peer feedback completed' (p.393). He adds that electronic systems such as WebCT can be used to post suggestions on the review process and thus save time.

In sum, peer reviewing is a complex skill and as numerous researchers have pointed out, students need to be guided through the process (Sluijsmans, 2002; Falchikov, 2005; van den Berg et al, 2006). It will only work as intended if all participants are engaged in the process. Important issues to consider in relation to student peer-review include how to prepare students, the types of activities to set, the methods of review to be used, and the groupings of reviewers and authors. This will depend largely on the particular needs and level of the students in the group (Hansen & Liu, 2005; Lunstrom & Baker, 2009).

2. Incorporating student peer review in teaching: Four case studies

Student peer review can be incorporated into curriculum in various ways. While the type and scope of the peer review exercise will depend largely on the aims and nature of particular subjects/courses and their specific disciplinary teaching and learning context, there are a number of (common) questions to consider when planning to implement student peer review. These include:

- What is the aim or main focus of the peer review exercise?
- What will the review process involve?
- How will students be grouped? In other words, how many reviews will each student receive and be required to provide?
- How will you facilitate and manage the peer review process? For example, how will you assist students to give useful feedback?
- Will the review be assessed? If so, how?
- How will you gather feedback from students on their experience of the peer review exercise?

This section of the guide presents examples of how academics from different discipline areas at the University of Melbourne implemented student peer review using PRAZE, an online system for managing student peer review (see Section 4 of this guide). Since its development in 2007, PRAZE has been used in more than 80 assignments in 40 subjects across eight faculties, with enrolments in individual subjects ranging from 15 to 230 students. These assignments reveal that peer review has been employed in diverse ways across the University. From teaching some of these subjects, and interacting with coordinators of others, we have gathered some insights into the use of peer review in these various settings and modes. The following four case studies capture and present some of those experiences.

Case A: Experimental Animal Behaviour

Academic staff: Raoul Mulder and Mark Elgar, Department of Zoology

Students: 60 (3rd year)

In *Experimental Animal Behaviour*, students work in groups of four to carry out a research project and write a report. Students choose a topic from a list of fifteen broadly pre-defined areas defined and then carry out their research under the supervision of a tutor. Although the research is carried out in groups, students prepare and submit individual reports.

Aims of peer review:

To help students improve the quality of their report before final submission

To introduce students to the type of peer review that takes place for journal article publications in Science

Process:

Raoul offered a 2-hour tutorial session in which he both described the peer-review process generally (in the context of the primary literature) and then gave more detailed descriptions of issues to consider when writing a review. After groups completed their research, an individual student would upload their individual draft report (which was not graded) for distribution to three reviewers: the group's tutor plus two student reviewers from different groups. Thus each student wrote two reviews for other students and received three (anonymous) reports on their own work. Report authors used the reviews to improve

their final submission, and wrote a 'letter to the editor' in which they detailed how they had dealt with the reviewers' comments.

Assessment:

Students' final report submissions were assessed together with their 'letters to the editor'. Students were also assessed on the quality of their reviews (10% for each review).

Student perceptions of peer review process:

Students responded very positively to the peer review process. A survey (approximately 80% response rate) showed that the majority of students (about 80%) believed the review process had helped them to improve the quality of their final submission. This is supported by written feedback comments such as "Reviews were a great help!", "Reviews were a very good aspect of this subject", "The peer review process was very helpful" and "I think the peer review exercise should be introduced into every Zoology subject".

The survey and comments indicated that most students had an extremely positive experience of peer review.

Lessons learned:

While no students disputed that peer review had been helpful to improving their work, a small proportion (about 20%) indicated that they were 'neutral' in their judgment about the peer review process. In particular, perceived variability in the quality of reviews can be cause for concern for some students, and this concern was indeed raised by a few students in *Experimental Animal Behaviour*. One possible response to this would be to increase the number of reviews each student received. Indeed, Cho *et al.* (2006) showed that the reliability of individual student reviewers tended to be modest and lower than that of instructors, whereas a collection of four to six peer reviewers produced very high levels of reliability. However, Raoul feels that the inclusion of a tutor as one of the (anonymous) reviewers most likely offered reassuring comfort to the majority. Variability in review quality is also a fact of life in any reviewing context and at any level. Appreciating this fact, and developing ways of dealing with this variation, are skills that he was keen to encourage as early as possible in a student's academic experience.

Case B: Peer review in Multimedia and Communications

Academic staff: Jon Pearce, Department of Information Systems

Students: 60 (2nd year)

Multimedia and Communications is a second year subject in which students engage in a web site development process with the aim of improving science communication through the use of multimedia. Students work in 14 groups of four or five students each. During seminar classes the students implement a user-centred web design process involving understanding a (fictitious) client's needs, and creating personas and scenarios to explore design requirements. They produce a prototype and then a final web site to deliver to their client.

Aims of peer review:

To provide students with formative feedback during their web design process

To provide feedback on each student's contribution to the group

Process:

Near the end of the web design process, when prototype sites were ready for some preliminary user feedback, each group submitted the URL of their site to PRAZE. These URLs were distributed to four or five other students in the class so that each individual received one web site to review and, consequently, each group received four or five student reviews of their work. Jon devoted one lecture

session to discussing peer review concepts and giving examples to guide students through the process. The review process was driven by an extensive online questionnaire and a standard web site usability checklist. PRAZE ensured that students did not receive their own group's site to review, nor indeed a site from their own seminar class (since they already had some familiarity with these sites). Two staff members also took part in the review process, reviewing all sites between them. Hence all sites received an additional review, resulting in each group receiving five or six anonymous user reviews, including one (unidentified) review from a staff member.

The comments in the reviews helped students to identify bugs and design issues, and to correct them before the final submission of their web site. They also had a chance to feed back their responses to the reviewers. The use of peer review in this subject differed from other subjects in that one of the questions on the review form asked reviewers to give a 'mark' to the work they were reviewing. The average of these marks was used as a small component of each group's final assessment (5%). This was an incentive for students to develop their website to as near completion as possible before submitting it for review.

At the end of the subject each student used PRAZE to peer-assess the performance of each member of the group (including themselves). Students were told that the rating derived from this 'team assessment' might be used to modify the final mark that each individual earned for their project. The review form asked students to enter an overall rating (from a drop down menu) as well as to write a brief comment on each team member. This approach was based on Oakley's paper on effective teamwork (Oakley et al., 2004).

Assessment:

Students' web sites were assessed as group projects with the students' reviews contributing 5 of the 35% of marks allocated for the project. The group self-assessment scores were used to help moderate the final individual marks for the project.

Student perceptions of the peer review process:

The 'team assessment' was a highly successful exercise in that, whilst most teams rated each other as making similar contributions to the project, in some cases issues were raised that showed workload or performance divisions within the team that needed attention. Sometimes one, or occasionally two, members of a group did not perform well – 'hitchhikers' or 'couch potatoes' in Oakley's terms – yet nearly always the whole group agreed with this perception, including the poor performing students themselves. In fact it was not uncommon for an individual to give themselves a lower rating than the rest of the team gave them. In one case, in a team of three students, two members rated every member with the highest rating ('A. Excellent') and wrote appropriately crafted comments to support each rating, yet one member rated himself as just 'C. Satisfactory' with a comment 'I believe I'm the most useless person in the team ... I believe its because my standard isn't that high compared to theirs. However, I really tried my best to come up with good ideas, participate, and do everything I could for the team.' Clearly it is important to look carefully at information gathered this way.

Responding to an end-of semester survey (about 50% response) all students stated that the comments they received in the peer reviews were 'helpful', with about three-quarters being positive in this response. About one half of the respondents considered that the peer review process actually helped them to improve the quality of their web site project while about one quarter thought that it did not. Students appreciated that the comments they received in the reviews came from a 'real world user' (i.e. a student, rather than a lecturer!), and commented on the value of the structured review form that helped to identify specific aspects of their work that they could improve. Only two or three students expressed some concern about the fairness of other students contributing to the assessment of their work (rather than the lecturer). Their concern was that other students might not review fairly due to either lack of understanding of how to grade the work, or due to deliberately trying to manipulate the

system either to be kind to others or to gain themselves a higher mark.

Lessons learned:

This second case study shows how peer review can be used as part of a *design* process rather than a *writing* process. It turned out to be extraordinarily valuable in this context. Not only did students see examples of the work of others – features of which often highlighted to them issues in their own design – but also they received multiple pieces of constructive feedback from peers identifying specific issues with their work, which they could address before their final submission. From a design perspective, this was a great way to give students the benefits of a formal, anonymous user review of their design work, without the students having the arduous task of setting up the review for themselves.

A few students expressed concern about fairness or potential to manipulate the system. Jon does not consider this to be of significant concern due to the anonymity of the system, and the fact that, where reviewers' scores contributed to assessment, each group's mark was an average of five or six independent reviews. In addition, the group peer self-assessment exercise was a valuable way of adjusting marks in cases where groups decided that there was significant unevenness in their contributions.

Case C: Reshaping Environments

Academic staff: Helena Bender, School of Land and Environments

Students: 350 (1st year)

Reshaping Environments is a fairly large first year subject with about 350 students. Students form groups to evaluate a physical site. Helena provides students with a list of sites on the University campus from which they choose one to evaluate.

Aim of peer review:

To help students improve the quality of a site diagram to be constructed by the group.

Process:

Each student was required to complete two anonymous reviews, which contributed towards their assessment for the task. Tutors participated in this review process so that each group received reviews from students in tutorials other than their own, as well as from their own tutor (although in subsequent iterations of the same subject Helena did not include the tutors). All three reviews were anonymous so the students were not aware of which review came from their tutor. They had time to respond to these reviews before submitting their final work, as well as give feedback to their reviewers. In the first year of using PRAZE in this subject, students repeated the process in their final assignment in which individual students submitted a poster to PRAZE (as a PowerPoint file) for review.

Helena ran a similar peer review process with students reviewing other students' mind maps. She also had them take part in evaluating each other's contribution to their group work using PRAZE. She surveyed the students before and after to assess their reactions to the process.

Assessment:

The assessment for the subject comprised both group and individual tasks. Students submitted a site diagram constructed by their group, individual PowerPoint posters and mind maps. Students were also assessed for their review of peers' work.

Student perceptions of the peer review process:

In the first iteration there were several technical teething problems with PRAZE that caused some frustration amongst the students. Nevertheless, students appreciated receiving the feedback on their diagrams. They found it challenging to decide what action to take when there were contradictory comments. Some felt that the effort and quality of the reviews varied too much, particularly for the

poster, and that they had been taken advantage of by their peers – there was a feeling that they had spent a great deal of time correcting someone else’s work that needed major revisions, while the person reviewing their work had provided few suggestions. Clearly there were some asymmetries in the effort contributed by the students.

The surveys in the second semester showed that the vast majority of students expected the process to be useful before they did a peer review, but about 20% found it less useful than anticipated. Just over a half regarded the writing and receiving of reviews to be of equal value to them and a similar number indicated that their written work improved as a result of the reviews received. Perceptions of their peers as qualified to provide critical feedback decreased from 70% to 58% after students had experienced the peer review process.

Lessons learned:

Helena’s experience highlights the importance of setting clear expectations to students before undertaking a peer review process. Students are put in the position of both reviewer and author and need to think through issues that they might confront in each role. For example, Helena suggests that students need to be encouraged to contemplate how it might feel to be in one of the following situations and how they might respond to the scenario:

- They are given a piece of work to review that is incomplete or of very poor quality.
- They are given a piece of work that is excellent – how do they offer constructive advice?
- They receive a review with comments that offer no guidance or useful criticism.

These reflections highlight the importance of devoting time to prepare students *before* a review process – and to providing ongoing support as students individually encounter such issues. To reflect on a piece of work and offer critical comments, identifying both strengths and weaknesses, is a skill students need support in developing. They benefit from being shown examples of reviews (both ‘good’ and ‘bad’) to help appreciate the kinds of comments that they could write. Helena also feels that they need to be encouraged to read such reviews from the perspective of ‘how would I feel if someone wrote this about *my* work?’ This applies equally to examples of good and bad reviews.

Case D: Programming Language Implementation

Academic staff: Harald Søndergaard, Department of Computer Science and Software Engineering

Students: 30 (3rd year)

Programming Language Implementation involves a project requiring students to work in small teams (up to four per team) to write a compiler for a small programming language. Harald uses student peer review as the second part of the three-stage project².

Aim of peer review:

To generate feedback on the computer code submitted by students in the first stage of the project and to allow students to learn from the other team’s solutions.

Process:

Students work in small teams (up to four) and submit their computer code to PRAZE and then individually review two submissions each. Hence each team receives, on average, five reviews. The reviews are marked by staff according to the quality of the review (allocated 0 – 2 marks per review, forming 16% of the assignment marks); but the reviews do *not* contribute to the authors’ marks. Students use an online review form to guide their review.

² For a detailed discussion of how student peer review was implemented in this subject, see Søndergaard (2009)

Assessment:

Students' group projects were assessed. Students were also assessed for their review of peers' work (4% of the subject mark).

Student perceptions of the peer review process:

Harald administered his own survey to gauge responses to the review process and also analysed the University's Quality of Teaching survey results. The overall feedback from students was extremely positive. On average they spent 2 hours on their reviews (ranging from 30 minutes to 6 hours!). They thought the double-blind process was appropriate and that the allocation of marks (16%) was just right. There was strong agreement that their peers were well qualified to carry out the reviews, equally strong agreement that the feedback was useful and that it was valuable to their own learning to see the work of others. They appreciated that not only did the process improve their own work but it also improved their skills to evaluate technical work. Students gained a better understanding of how their group was progressing, and reported a 'feeling of being part of a learning community'.

Lessons learned:

Harald had a philosophy of driving students' learning via a challenging small group project, completed in three stages, with peer reviewing as the middle stage. He deliberately avoided any use of 'students grading students' in order to help encourage a sense of 'learning community' amongst the students. This was also supported with online discussions. This has been a highly successful endeavour – both from Harald's perspective and from that of the students. He has observed a high feedback score on the Quality of Teaching survey, which indicates that students do see peer review as a valid form of feedback. Harald's experiences endorsed his approach of grading the quality of students' reviews himself, whilst limiting student involvement to providing helpful formative feedback to their peers.

Reflection on the case studies

These case studies highlight the variety of ways in which peer reviewing can be implemented in teaching settings across a range of year levels from first year to third year. These illustrate the variety of submission types (essays; links to web pages; diagrams; presentations; mind maps; computer code), the nature of the reviews (critiquing written work; informing design; critiquing diagrammatic representations; commenting on computer code), as well as the approach to the reviewing process (who takes part in the review; whether group work is submitted by the group or by individuals; which parts, if any, of the review process contribute to assessment).

In each of these four cases peer reviewing was part of a formative process to help students improve their work. In no case was it used solely at the end of a project or assignment to provide summative feedback. There was a mixture of ways in which the reviews contributed to the students' assessment. In all cases the process was well received by students and generally recognised as a means through which they could improve the quality of their work.

An important theme to come from this is the crucial role of briefing the students on how to carry out a review and on what constitutes a good review. This was critical to the success in each of the cases described. Recognising that this was a novel experience for most students, staff set aside class time to introduce the idea of peer review to students and to provide guidance as to how to go about the process.

3. Reflections on implementing student peer review

This section of the guide presents a reflective discussion of two academics' experiences in facilitating and implementing student peer review. After a brief description of the reasons for using student peer review in teaching (3.1), we examine students' perceptions of the experience (3.2) and discuss some issues for consideration in implementing student peer review (3.3).

3.1 Why use student peer review in teaching?

Our interest in student peer review was motivated by the fact that we wanted to provide students with feedback that promotes a genuinely reflective cycle of learning. In a university environment, the typical source of 'feedback' remains some form of commentary accompanying a mark attached to the final version of a submitted assignment. The great limitations of this time-honoured approach are that only a single perspective is provided (the teacher's), and that there is usually no further opportunity for the student to improve on the assignment. It is not surprising then, that students have little incentive to reflect on, or learn from, this 'feedback'. The obvious solution to this problem is to provide students with repeated cycles of feedback, but of course these place considerable demands on teaching staff, especially in subjects with high enrolments.

We have found that student peer review of draft work before submission, combined with the development of online tools to facilitate its administration, provides an effective and potentially widely-applicable solution to the challenge of providing useful feedback. Although not all software packages are similarly able to accommodate variation in formats, in principle there are no restrictions on what can be reviewed (uploaded documents, websites, programming language, designs, events, art, or buildings, to name just a few), or in how formally or informally the process might work (structured versus open, online discussion, face-to-face, double-blind anonymous). The case studies in Section 2 introduce some of the possibilities of formative peer review, demonstrating that peer review can work in diverse contexts.

Although most students participating in peer review have had little or no previous exposure to it in their studies, evaluations suggest that their experience is overwhelmingly positive, and that they learn both from providing and receiving reviews. Many students comment that completing a structured, comprehensive review for someone else is an illuminating way of becoming aware of areas that require attention in their own work. They gain an appreciation for the fact that reviewers can differ widely in opinion and competence (or can have surprisingly similar, independently derived opinions) and learn how to distinguish between helpful and unhelpful forms of feedback. The exercise has discipline-specific learning benefits, but it also serves as an introduction to a fundamental generic skill: providing and responding to constructive critical feedback. Most importantly, it allows students the opportunity to benefit from peer input and to submit work for final assessment that genuinely reflects their capabilities and talents.

Peer reviewing also offers an experience that students often comment as lacking in their university studies: the opportunity to see the work of other students. As assessors, we are accustomed to viewing many dozens of assignments on a given topic and tend to forget that students submit their assignments in relative isolation. They are not privy to the same perspective and often have no way of knowing how the standard of their work compares to that of others.

3.2 Student perceptions of the peer review experience

An important factor in evaluating and reflecting on our approaches to teaching and learning involved seeking feedback from students about their learning expectations and experiences. To examine students' perceptions about the effectiveness of the peer review exercise, we asked teaching staff to administer a short survey before and after students participated in the peer review process. We report here on the aggregation of survey responses from four subjects for which we have complete before and after data. We will not report detailed differences between these subjects except where there is an occurrence of particular interest or significance.

Survey methodology

The set of four subjects comprised:

- a large first-year subject (about 150 students)
- two third-year subjects (about 20 and 80 students respectively)
- a small post-graduate subject (about 16 students)

The peer review exercises in these subjects involved both group and individual work in the submission of essays and design projects. In each case the peer reviews were used to provide formative feedback to the students.

The four questions on the pre-survey were:

1. I had previously participated in peer review before undertaking this subject.
(yes, more than once; yes, once; no; not sure)
2. As a learning tool, I expect peer review will be:
(very useful, somewhat useful, no opinion, not very useful, useless)
3. I expect I will learn most from:
(writing reviews of other students' work; receiving reviews of my own work; writing & receiving reviews; not sure)
4. I think that my peers are well qualified to provide me with critical feedback on my work:
(strongly agree; agree; neutral; disagree; strongly disagree)

After the students had completed their peer review experience, we administered the post-survey with the following questions:

1. As a learning tool, peer review was:
(very useful; somewhat useful; no opinion; not very useful; useless)
2. I learnt most from:
(writing reviews of other students' work; receiving reviews of my own work; writing & receiving reviews; not sure)
3. I thought that my peers did a good job in providing me with critical feedback on my work:
(strongly agree; agree; neutral; disagree; strongly disagree)
4. I think that I improved my written work as a result of the reviews that I received or wrote:
(strongly agree; agree; neutral; disagree; strongly disagree)

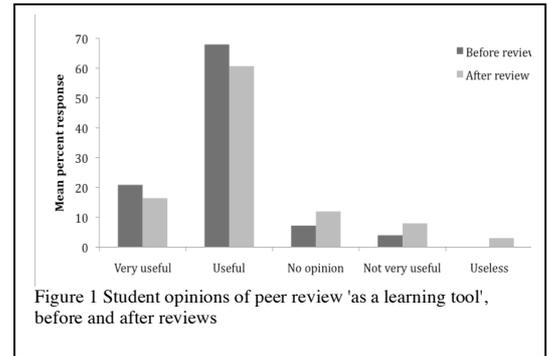
For each question in the pre-survey, the response rate of over 95%, and in the post-survey over 90%. Hence the before/after data presented in the next section relates to 278 and 199 students respectively.

Over half of the students surveyed had experienced a peer review process previously in one form or another. This proportion was particularly high for the two third year subjects. Surprisingly, students in the post-graduate subject indicated the least experience with peer review.

Survey Results

Do experiences with peer review match expectations?

Students had high expectations of peer review being of value to their learning, with over 90 per cent expecting either 'very useful' or 'useful' outcomes. This was remarkably uniform across each subject. After using the peer review system, students' post-user evaluations closely matched their earlier expectations, but with a slightly reduced 78 per cent identifying peer review as 'useful'. These before-after comparisons are shown in Figure 1.

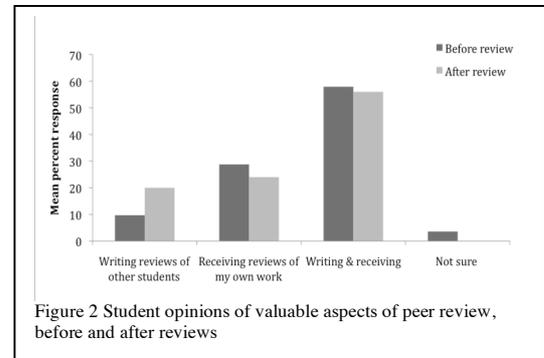


Which aspects of the review process do students find most useful?

While many students expected receiving reviews of their own work to be the most valuable outcome of the review process (30%), the majority (60%) showed awareness that both the receiving and writing of reviews would be of value (the post-graduate was the most extreme with nearly all students indicating 'writing and receiving' as their most useful expectation of the process). Across all subjects, receiving reviews was expected to be more valuable than writing them, with this being a particularly strong feeling in the smaller third-year subject.

After the review process students' perceptions showed a more equal balance between receiving and writing reviews across the individual subjects, except the third-year subject, which maintained an atypical bias towards receiving reviews. This is an interesting outcome in that it suggests that, while many students expected the value of the process to be one of receiving reviews of their own work, quite a few students shifted this opinion and considered writing reviews to be of more value than receiving them (Figure 2). This was highlighted by the small post-graduate group that moved away from having very popular prior expectations of the value of both writing and receiving reviews, and indicated more value in the writing process, and to a lesser extent receiving reviews.

'I found it a valuable process, which also helped me to look at my own work more critically. I think peer review is an important skill to be able to learn, and I think it was introduced well in this subject.' (student)



'Feedback on my reports were extremely helpful as each reviewer saw mistakes that neither I nor other reviewers saw, which helped me in improving my own report. I also enjoyed writing reviews for other people's reports, as it helped me reiterate what was right and wrong to do in my own report.' (student)

How do students regard their reviewers?

When students were asked before the peer review process to rate how qualified their peers were to provide critical feedback, about two-thirds 'agreed' or 'strongly agreed' that their peers were qualified. However, about 27 per cent were unsure or had no opinion (Figure 3). After carrying out their reviews, there was a slight shift down in confidence. Much of this was due to the average score on this question falling for the large first-year class – quite a few (17%) indicated 'disagree' or 'strongly disagree' in response to the question.

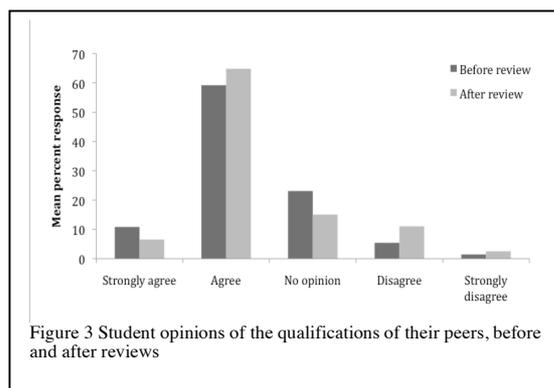


Figure 3 Student opinions of the qualifications of their peers, before and after reviews

Did peer review improve their written work?

Students were finally asked whether they thought peer review had aided in improving their written assignments. Again students were overwhelmingly positive: over 71 per cent 'agreed' or 'strongly agreed' that peer-review had helped improve their work (Figure 4)

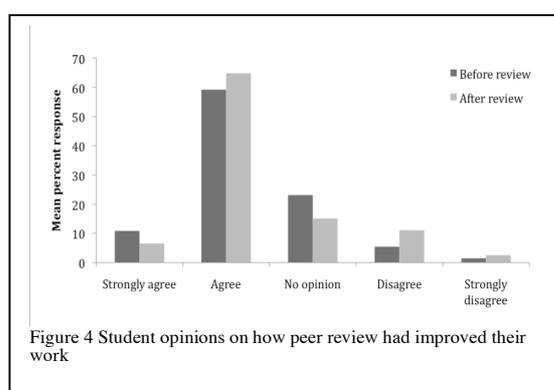


Figure 4 Student opinions on how peer review had improved their work

Student comments

Students were given the opportunity to write comments both before and after the review process. About a third took up this opportunity each time. The main concerns expressed before reviewing (by about 30 of the 85 respondents) related to concerns about the likely quality of reviews – that they might be written by students who are not able to give good advice and might not put in the effort to provide high quality feedback. There were also some concerns expressed that the reviews might form part of a mark for the subject and that students might give better marks to their friends (all reviews were anonymous and none in these subjects was used for summative assessment).

These concerns were reflected a little in the post comments, but not strongly. About 10 comments indicated a concern about the quality of the reviews but the majority of comments (70 percent) were positive. Whilst most of these positive reviews made general comments about helping to improve their writing, some in addition expressed concerns about the difficulty in relying on other students to give high quality reviews. Four students specifically mentioned that receiving reviews was not helpful, but writing them was helpful.

'The reviews on my work helped however they were a bit confusing as they all said different things. One review would say something is good but another would say that it needs a lot of work, so it made it harder to pin point exactly what was wrong with my article.' (student)

'Although it took time to do, it made you look at your own work in a different view, as well as exposing you to alternate views of others. Good system.' (student)

These general comments support the impressions gained from the survey in that there was some apprehension before the review process as to the ability of students to write useful reviews. The post review survey and comments suggest that this apprehension was largely unwarranted and that there was some recognition that much of the value in the process came from actually writing reviews of others' work.

Although student peer review has been trialled in relatively a small number of subjects across the university – and its effectiveness evaluated in even fewer – these evaluations suggest that it is generally perceived to be effective in improving learning, and is well received by students. However, like any other aspect of teaching, student peer review is not a one-size-fits-all solution and staff should be mindful of potential problems and pitfalls. In the following section, we address academic and practical issues associated with the administration of student peer review, highlight common concerns raised by students and staff, and speculate about future prospects for student peer review.

3.3 Key considerations in implementing student peer review

While our experiences have been extremely encouraging, student peer review is not without its difficulties. A recurring theme emerging in student evaluations is concern about variation in review quality. Below, we consider this and some other common concerns raised by both instructor staff and students in relation to student peer review (disparities between peer feedback and assessment outcomes, plagiarism, and variable participation), and offer our perspective on how some of these concerns might be addressed.

a. Dealing with variation in review quality

Student peer reviews can vary dramatically in their quality, namely the degree to which they provide feedback that is accurate, specific and helpful. The two primary contributors to this variation are likely to be either lack of skill and experience, or lack of effort and motivation. Improved reviewer training and calibration can be effective ways of improving reviewer competence to at least a minimum standard, and structuring of review forms can help focus the attention of reviewers on pertinent issues. Summative assessment of reviews is often employed to improve the motivation of reviewers to perform their task to a high standard. The impact of variation in review quality on individual students can also be reduced by either increasing the number of reviewers, or the diversity of reviewers (e.g. by including 'expert' reviewers as well as student peers in the review process). Finally, we imagine that future online peer review systems could be designed to reveal anonymous information to authors about reviewers (for instance, scores they might have obtained from a calibration process), which might help an author determine whether feedback emanates from a relatively reputable or uninformed source.

Reviewer training, calibration, and structuring of reviews

Most students have never participated in any kind of academic peer review before, and therefore advice and training in reviewing are critically important forms of support that should accompany peer review assignments. For instance, in our *Animal Behaviour* and *Experimental Animal Behaviour* subjects we offer a 2-hour tutorial session in which we both describe the peer-review process generally (in the context of its role in academic publishing), give detailed descriptions of issues to consider when writing a review, and answer questions about the review process. The session includes advice on the process of reviewing (reading and annotating a manuscript, preparing and proof-reading the review, completing the review form), the importance of providing both praise for the positive aspects of a study as well as highlighting the weaknesses, examples of helpful and unhelpful reviews, tips for time management of reviews, and references to online resources. Students nevertheless often report that they find their first experience with peer review overwhelming. Thus, while tutorials are important and necessary, they may be insufficient on their own without some form of opportunity for 'learning by doing'.

Some online peer review programs provide students with the opportunity for 'practice' by including a calibration step (e.g. Chapman 2001). Undertaking calibration allows students to test and develop their reviewing skills on set pieces of work of varying quality, comparing their review answers to a 'correct' reference set. They can only graduate to peer review after they have demonstrated competence at the calibration task by passing it. While calibrations are an excellent way of assisting students to learn what to look for in reviews, it seems likely that not all assignments will lend themselves well to a calibration step (for instance, if there are many or nuanced forms of 'correct' answers). Thus, calibration will be better suited to some forms of assignments than others.

An important and effective way of lessening variation in review quality is through careful design of review forms. These can be structured to require reviewers to answer "yes" or "no" to key questions, rank elements along a Likert scale, indicate the presence or absence of information using checkboxes, or provide feedback in an open format such as a text box. If the preamble instruction to a text box is carefully worded, this can also help the student being reviewed to judge the value and authority of the comments. For example, if the reviewer is asked to "Justify your earlier comments by referring to principles developed during the subject", this should prompt a justification that might help the recipient judge whether or not the reviewer shows insight into the issues. Structured forms are not only very helpful in guiding reviewers with cues to the issues that need to be addressed, but in our experience they also provide an excellent way to reinforce attention to the criteria by which the work will eventually be assessed.

Assessment of reviews

Many, but by no means all, subjects employing formative student peer review include assessment of the reviews as a component of a student's overall grade for the subject. The rationale is that making the reviews assessable will motivate greater investment, and thus elevate the mean quality of reviews. While few instructors would doubt that this logic is reasonable, there are some problems associated with summative assessment of reviews. First, it is unclear how important motivation is (relative to lack of skill, for instance) as a contributor to variance in review quality. In classes where motivation is already high, assessment may do little or nothing to flatten out variation. Second, making reviews assessable can substantially increase the assessment burden on staff. This is particularly problematic because it is often desirable to have multiple reviews for each assessable item (see below); each additional review adds one class's worth of additional assessment. Finally, evaluation of reviews is not straightforward; to gain an informed opinion of the merits of a review, the assessor will often need to read not only the review but also the work being reviewed. Depending on the volume of the assessment item being reviewed, this could rapidly mushroom into an unmanageable assessment load. Thus, the question of whether reviews should or should not be assessed may depend on such factors as whether alternative means of motivating investment are desirable or available, how many staff and other resources are available to a subject coordinator, and to what degree the subject administrator and staff are willing to invest in assessment.

Given these issues, if assessment is primarily intended as a driver of motivation, there may be alternative means of achieving this end without assessment. Possibilities might include 'award'-based or other incentives, or promotion of competition and individual status through publication of reviewer rankings. To our knowledge, few of these have been trialled or evaluated in terms of their efficacy.

If assessment is deemed necessary and useful, the fairest and most effective solution to many of the problems above may be for scoring or assessment of reviews to be done by the student recipients of those reviews. Review recipients are presumably well-placed to comment on whether or not a particular review was helpful, particularly because they are able to make comparisons with other reviews of their work. While there is some potential for bias (e.g. authors rewarding flattering rather than insightful and critical reviews), such bias could be mitigated in a number of ways. These include careful design of the

evaluation forms to emphasise the desired criteria, and mathematical adjustment of mean scores for reviewers to eliminate outlying low or high scores.

Increasing the number of reviewers

Peer review tends to be most effective when multiple reviewers are involved. Reciprocal review involving only a single reviewer is problematic for a number of reasons. From the perspective of the reviewer, a single peer review exercise gives limited opportunity to practice a task that is challenging and normally requires repeated practice to be done skilfully. From the perspective of the author, a single review offers no opportunity for learning by comparison because each student is exposed only to a single opinion. This tends to reinforce simplistic notions about whether or not particular approaches, ideas or forms of content are 'valid' or 'correct', rather than the more subtle and complex reality that reviewers vary in opinion. Authors are furthermore not challenged to apply their learning to filter among reviews and rationalise decisions they make about whether or not they adopt or reject reviewer suggestions (for instance by juxtaposing or emphasising 'majority' views).

Perhaps most importantly, the potential for inequity is large with a single reviewer, because variance in review quality is spread across individuals in the class – certain students may receive only one poor review, while others receive an excellent review. If students have limited experience with the review task, there is a high likelihood that some reviews will be poor. In one of our subjects *Animal Behaviour*, students received only a single peer review of their work in 2007, whereas those in the 2008 cohort received two reviews, and those in 2009 three reviews. Student responses in 2007 to the statement "Peer reviews helped me to improve my article" ranged from 1 (strongly disagree) to 5 (strongly agree), with a mean of 3.3 ± 1.3 S.D. ($n=88$), while in 2008 and 2009 the means increased and standard deviations decreased to 3.5 ± 1.1 ($n=80$) and 3.8 ± 1.1 , respectively. It is difficult to imagine which possible confounds across cohorts within this subject might offer alternative explanations for this significant trend, since every other aspect of the process remained identical. Thus, for this subject at least, there is evidence that overall satisfaction with the process increased, and variance around the mean decreased when students received more reviews. Cho *et al.* (2006) have also shown that the reliability of individual student reviewers tends to be modest and is lower than that of instructors, whereas a collection of four to six peers produced very high levels of reliability. Thus, as a general rule of thumb, instructors should consider a minimum of two to three reviewers per assignment, and ideally more (e.g. four to six).

Increasing reviewer diversity

Although increasing the number of reviews is an effective way to dampen variance in review quality, this can also be achieved in other ways. For example, 'expert' reviewers as well as students can be involved in the peer review process. In *Experimental Animal Behaviour*, students receive three reviews, one of which is by an instructor. In principle, this guarantees every student of at least one review by a reviewer with some experience and who was knowledgeable about their subject. Notably, in this subject students showed very high levels of satisfaction with the peer review process.

Having instructors involved in writing reviews can also have a beneficial moderating influence on the review process. In many review systems authors can send (anonymous) feedback to their reviewers to comment on, or thank them, for their reviews. Knowing that one of these feedback comments might be received by an instructor minimises the chance of any malicious grievances being sent to a critical reviewer!

We conclude this section by noting that while for reasons of equity it is preferable to reduce variation in review quality, it is important that this not be confused with reviewer variation in emphasis or opinion, which we consider to be an appropriate, authentic and indeed pedagogically vital facet of peer review. Exposure to differences of reviewer opinion can initially be quite confusing and frustrating for students.

However, we firmly believe that it both prepares students for professional reality (particularly in publishing!) and encourages them to develop means of assessing, organizing and rationalizing their response to such variation.

b. Disparities between reviewer feedback and assessment outcomes

Sometimes, students who are initially enthusiastic about peer review become disappointed or disillusioned about the process when their expectations about grades for their work are not matched by the outcome. Typically, this results because the student has received highly positive reviews on the draft assignment from several reviewers, but then receives a relatively low final mark. The disparity between the assessor and the reviewers in assessment highlight the fact that in some cases, the skills expected of reviewers may exceed those that have actually been attained (i.e. they failed to identify important problems with the draft assignment). This could be for a number of reasons, but most likely insufficient training and experience. Increasing time dedicated to training, introducing a calibration step, and improving review rubrics are some ways in which the likelihood of this sort of (fortunately rare) negative experience can be reduced.

However, it is perhaps equally important to remind students that the opinions of peer reviewers are always potentially limited, and that it would be unwise to interpret them as predictors of how their work will fare when assessed by staff. In the end, these reviews are simply the considered opinions of several peers – not the analysis of a professional assessor. Hence students should also be reminded that part of the aim of the exercise is for them to gain skills in judging the opinions of other people.

c. Plagiarism

An oft-cited concern around peer review is the possibility that reviewers might plagiarise ideas or other material from the work they review. While we do not doubt that this concern may be legitimate, we have relatively little experience with this problem and therefore can offer only limited suggestions here. In general, we have found that there are several effective ways to reduce the likelihood of plagiarism, at least some of which should be readily transferrable to other teaching contexts. Plagiarism is most likely when all students share the same assignment topic. Thus, a simple way to reduce the potential for plagiarism is to set multiple topics within a class and allocate reviews reciprocally such that no reviewer receives an assignment to review that is on their own topic. To ensure that any such allocation is mathematically possible, assignment topics should ideally be quota-restricted; this ensures an even distribution of students across topics. Note however, that at present, only one software package offers the functionality of applying rules to the distribution of reviews among students (PRAZE; Table 1). For subjects in which plagiarism is a significant risk, the *Turnitin* platform offers the possibility of peer review and plagiarism detection within the same assignment.

In some cases, students fail to submit work for review, or fail to provide a timely review on submitted work. This can create problems for the administrator of the review process because any delay disadvantages the student peer awaiting work for review or from a reviewer. One (perhaps Draconian) way to motivate compliance with deadlines is to make these hurdle requirements for passing the subject; failure to do otherwise leads to an automatic failure. Nevertheless, this solution is not appropriate when students are unable to participate because of circumstances beyond their control (e.g. illness). Various solutions are possible, but these are generally impractical if the number of defectors in a class is very high. For instance, one peer review software package (PRAZE, Section 4, Table 1) allows reviews to be distributed only among students that have submitted work, reducing the problem of 'empty' reviews or reviewers. Nevertheless, any late reviews, with genuine justification, still need to be handled manually.

While the software packages that exist to manage peer review at present (see Section 4) do not yet offer elegant solutions to this problem, we imagine that there are interesting ways in which this could be

addressed in the future. One possibility would be to allow students to 'bid' online for the opportunity to review late or unallocated work for additional credit.

d. The role of group self-assessment

A slightly different form of peer review offered by one system (PRAZE, Table 1) is the ability for students working within groups to self-assess the performance of each member of their group (including themselves). Although the other peer review systems reviewed in this booklet do not offer this feature, there are other standalone systems that do (for example, *iPeer* from the Centre of Instructional Support in the University of British Columbia). A common issue with group project work is the equal allocation of assessment to all members of the group when the contribution to the project might vary considerably amongst the group members. This does not necessarily mean that students should receive differentiated marks; there are situations where one might want to make it a responsibility of the group to ensure that all members contribute equally and not interfere in this process. However, in spite of assignment 'cover sheets' soliciting a weighting from each student as to the amount of work they contributed to the project, it is often difficult to obtain reliable feedback from students as to how marks should be allocated.

The group self-assessment process within PRAZE allows each group member to complete an online form in which they rate each group member, including themselves, on a performance scale and enter a comment about each student's contribution. The output of this process is a table listing each member's average score relative to that of the group, together with all the comments written about each person. This can be used to weight individual's assessment scores if necessary.

Our experience is that students are generally refreshingly honest in this process. If there are one or two students (from a group of five or six) who are not contributing fairly to the group then the whole group, including the offenders themselves, tend to acknowledge this fact in their ratings and comments. Whilst one might not want to blindly apply a weighting to the students' scores, this process provides an easy way to sort for extreme offenders, read the groups comments, and make a considered judgment. Oakley et al. (2004) offer sound advice as to how to apply such weightings.

e. Open questions

While some research has been carried out into effective designs of peer feedback activities (see, for example, the study by van den Berg et al., 2006), in general there are limited evaluative data on student peer review. We still know relatively little about the effectiveness of peer review, and thus there are many outstanding questions that lack satisfying answers. Are first year students too focussed on a 'delivery' mode of education to accept feedback from their peers? Are students' attitudes to peer review affected by their discipline or cultural background? Is feedback on writing exercises perceived to be any more or less effective than feedback on design exercises or oral presentations? For many of these questions, student surveys carried out as a part of the peer review process provide a relatively easy and straightforward means of evaluating aspects of the peer review process. We therefore encourage academics to consider administering such surveys wherever possible.

We have used the peer review process mainly in the context of providing formative feedback. However, there is also significant potential for students to participate in summative feedback (i.e. provision of grades, see for example 'Assessment of reviews' above). In our subjects we have generally linked peer review to assessment in two ways: a tutor assessing the quality of the reviews, and students assessing the quality of the work they were reviewing. We need to explore other modes of incorporating assessment and how effective they might be. While there are understandable concerns about the potential effects of student inexperience and potential for bias, Cho *et al.* (2006) point out that instructor reviews may similarly be subject to issues of reliability and bias, and demonstrate convincingly that under appropriate conditions, peer-generated grades are sufficiently reliable and valid to be used widely

in university settings. Indeed, it appears that negative student *perceptions* about the validity and reliability of peer-generated grades are a more significant impediment to their implementation than the validity and reliability of the scores themselves.

Because peer review is not yet a common feature of university curricula across campus, part of what has excited student interest in our subjects may be the novelty of their experience. If more and more subjects experiment with peer review, is it possible that the effectiveness of peer review will diminish, for instance because students tire of using it as a learning tool? While we cannot predict the longer-term acceptance of peer review, we do not believe their current enthusiasm is solely due to its novelty. If peer review is employed in a way that helps students learn, and if that learning is rewarded, we believe that peer review will become increasingly commonplace in universities.

The following section of the guide discusses how student peer review can be organised and managed using online tools.

4. Managing student peer review using online tools

Given that the pedagogical merits of formative peer review are so well established, it is perhaps surprising that student peer review is not a more pervasive feature of university curricula. One reason is that administrating anonymous peer review without the aid of custom-designed software is so onerous that it remains a potent disincentive to implementation, especially when classes are large. Online tools promise to significantly reduce this burden, and are therefore an important part of the peer review landscape.

Most academics will already have encountered online submission and review in the course of publishing their research. The process of managing student peer review is conceptually identical, but the range of online tools available for management of peer review in an educational context is surprisingly limited. Furthermore, the platforms that are available differ widely in their capabilities and the degree to which they are customisable or suitable for the diversity of academic needs.

In this section we review currently available software for formative peer review, and summarise the capabilities of different packages. This summary is not intended to be comprehensive and we may have omitted certain software packages with which others are familiar. Our intention is to highlight well-established and accessible platforms (for which good documentation exists), and to highlight their capabilities, strengths and weaknesses. We also describe in some detail the University of Melbourne's in-house program PRAZE. Programs which have, as their sole purpose, the facilitation of summative assessment (group work and/or grade adjustment) such as *iPeer*, *WebPA* and *SPARK*, are not discussed here.

Contexts and software capabilities

Peer review is used in varied contexts, and some software packages are better tailored to particular needs than others. Table 1 summarises key features of the five most widely-used programs for peer review.

The most common requirements of peer review are the ability for students to upload work for review, an automated review distribution system, and an online review form. These features are offered by all the programs listed in Table 1. Beyond these simple elements, particular software systems offer additional features, such as the ability to:

- include non-uploadable work for review (eg artwork, URLs, oral presentations)
- apply rules to distribution (ie exclude particular reviewer-author pairings), or manually modify them
- tailor review forms to the requirements of a specific assignment
- enable students to learn by accessing reviews of the same work by others
- accommodate structuring of the class within a review assignment
- enable reciprocal feedback on performance among members of a group
- enable authors to send feedback to the reviewers of their work

In practice, staff typically seek to do one or both of the following:

- a) have students in a class provide each other with feedback or commentary on a draft version of individual or group work,
- b) have members of a group evaluate each others' performance.

However, within these broad classes of use, there is considerable variation (Section 2 highlights some of this diversity using case studies from the University of Melbourne). Key differences across subjects relate to such parameters as:

- whether students work individually or in groups;
- whether groups submit individual or collective work for review;
- how many reviews each student is expected to complete;
- whether staff as well as students are involved as reviewers; and
- how reviewer-author pairings are constrained by topic choice or class grouping.

The best choice of software for a particular subject or assignment will depend on the level of individual customisation that is required (i.e. the degree to which these differences can be accommodated; sometimes the requirements of the assignment match commonly-used or existing formats), and the administrator's level of experience and/or willingness to experiment (technophobe or digital native?). Below, we briefly describe each of the web-based reciprocal peer review systems that are listed in Table 1.

CPR (Calibrated Peer Review)

CPR (Russell 2004) differs from most other peer review software in its strong emphasis on training or 'calibration' of reviewers in the skills of reviewing. A CPR assignment consists of three stages. In the first *text entry* stage, a student explores online material on an assignment topic and submits text based on this material in response to a question. During the subsequent *calibration and review stage*, students evaluate example texts, also known as calibration essays, by answering questions about them. The students' responses are compared to a reference set of correct answers, and the student is able to access these for purposes of feedback and comparison. Until the student attains a minimum standard of competence in evaluation (measured as a percentage of 'correct' answers against the rubric), they are unable to graduate to reviewing the work of their peers. Once they have mastered the calibration, they are permitted to evaluate texts submitted by peers, and their own text entry. In the *final stage*, each student receives feedback both on the quality of their reviews, as well as a summary of evaluations by other students of their text.

The emphasis on reviewer calibration is beneficial both for reducing variance among reviewers in their proficiency and, in the process, fostering confidence among authors in the expertise of their reviewers. On the other hand, calibration may not be suited to all assignments (e.g. where there is no 'correct' answer), and might also discourage creativity in some contexts. Structuring within a class is not possible in CPR, so review of peer performance in any context other than written work is excluded. This tool is best suited to content learning and coaching of students unfamiliar with peer review, by means of iterative exposure to key criteria.

PRAZE (Peer Review from A to Z for Education)

PRAZE (Mulder & Pearce 2007) is a sophisticated online system that facilitates flexible management of all aspects of peer review. It allows staff to set up, customise and manage a peer review process within a subject, so that students can then anonymously review each others' work, send and receive feedback on their work, and/or carry out a peer self-review of group work. Students can complete online sign-up to assignment topics, create their own, or be allocated to topics by administrators. In principle, there is no limit on what can be reviewed (traditional written assignments, files of any format, external works or events such as URLs, artwork or oral presentations). Administrators can specify rules that govern the distribution of work (e.g. to mandate or eliminate the possibility of reciprocal review among group members), and can enable involvement by both peers and teachers in the review process. After the

review process, authors can complete a feedback form that is returned to their reviewer. Review and feedback forms can be easily designed from a menu of drag-and-drop elements, or drawn and modified from a library of existing forms. Assignment management is facilitated by integrated summary tables. A desktop widget provides updates on the status of the review process.

The main strengths of the PRAZE platform are its uniquely flexible and powerful customisation features. To the best of our knowledge, it is the only peer review program currently available that is capable of managing any type of peer review. Staff often seek to employ more than one stage and type of review assignment within a unit of study (e.g. peer review of a group-generated document, followed by within-group assessment of member performance), and management of this process within a single system is easier and more elegant than employing several different programs. At present, the main drawback of the system is that it is not widely available. It is currently restricted to staff of the University of Melbourne, but a standalone version and one that will integrate with the Sakai platform are currently under development.

SPA (Self and Peer Assessment) in Blackboard

The ability to conduct peer-review is a relatively new feature of Blackboard v8.0, offered through the 'Self and Peer Assessment' (SPA) Subject Tool. The assignment set by the subject administrator consists of one or more questions (which may be simple or complex), each of which has one or more criteria associated with it. In the first phase of the assignment, each student submits a response to the question(s). This is followed by an evaluation period during which the same students review their own and other submissions and assess them against the criteria. An optional model response can be provided during the evaluation period to help reviewers. Submissions and evaluations can be monitored, and bundled into a ZIP file for export or import.

The most attractive feature of the Blackboard SPA tool is its integration with student lists and other elements of a subject's LMS online presence. However, at present this tool has many limitations which restrict its utility for peer review. Structuring within a class is not possible – all students must work on the same topic. Every student account is automatically included in the pool for review distribution (even if a student has not submitted work, or the account is a 'student view' account created purely to allow staff to simulate the experience of viewing content from a student perspective). As a consequence, many students will be allocated 'empty' submissions or reviewers. Finally, work cannot be submitted late or by proxy through the subject administrator, even if the student has a legitimate reason for delay. In its present form, this tool is probably best suited to whole-class content-learning-based assignments.

SWoRD (Scaffolded Writing and Reviewing in the Discipline)

The primary goal of the review process implemented by *SWoRD* (Cho & Schunn 2007) is to improve student writing skills by means of repeated review and feedback cycles. Students learn to assess written work according to key attributes: *flow*, *argument logic*, and *insight* (before embarking on reviewing, students can participate in a calibration exercise, but this is optional). Students (authors) first submit written work which is distributed to five or six peers (reviewers), who grade the paper according to the three key attributes and offer advice on how to improve it. The author then has an opportunity to revise the paper, and it is sent back to the same reviewers for final review. Authors rate reviews on their *accuracy* and *helpfulness*. Composite ratings for each reviewer from the five or six authors are computed to determine a grade, which provides an incentive for reviewers to invest in the task. The rubric can be changed to include criteria other than the three key default attributes, but criteria are limited to three dimensions.

SWoRD is particularly well suited to assignments in which the development of writing skills is critical. It is less well-suited to content learning via peer review than some other systems (e.g. PSA and CPR), and its restrictive design for criteria means that it is not ideal for assignments that do not involve writing

per se. Like most other systems, it does not accommodate group structure, and work can only be randomly distributed among students. At the time of writing, the program was being extensively revised. It is not yet clear, however, whether this will involve new functionality.

Turnitin

The *Turnitin* system is best-known for its plagiarism-detection tool. More recently, peer review functionality has also been enabled for *Turnitin* assignments. Staff may upload their own assignment or select one from an online library. Students then submit papers on the assignment topic (the format of uploaded work is currently restricted to MS Word documents). Reviewers are assigned to papers in one of three ways: automatically (random), manually by the subject administrator, or open (students select papers for review that interest them). Reviews are completed in an interface that allows the review rubric and document under review to be simultaneously viewed. The text being reviewed can be marked and referenced in the review, enabling students to match rubric comments in the review with the relevant point in their document.

The *Turnitin* peer review system is particularly attractive for subjects in which plagiarism detection matters (and which already employ *Turnitin* assignments), and for which review material is submitted in MS Word document. The library of pre-written review assignments will also appeal to instructors who may be new to online peer review. However, the software is not designed to deal with non-standard assignment formats, or those in which students are allocated to different topics within a particular assignment.

Table 1. Summary of commonly-used web-based software packages for electronic management of formative student peer review

	CPR	PRAZE	SPA (Blackboard)	SWoRD	Turnitin
Primary use/niche	Peer review tasks with emphasis on training of reviewers	All aspects of peer review, especially those involving groups	Simple peer review tasks where admin level control is unimportant	Assignments with emphasis on improving writing skills	Class-based peer review where plagiarism detection is also desired
Description	Peer and self-assessment with explicit development of reviewing skills through calibration and repetition	Highly flexible and customisable peer review system designed for any review context	Peer and self-assessment module available as building block	Web-based reciprocal system emphasising development of writing skills	Anonymous or attributed peer review linked to Turnitin assignments
Main advantages	Ability to calibrate reviewers before they engage in peer review; maturation of understanding of criteria; online tutorial	Customisability of all aspects of the review process; works from within LMS with user recognition	Integrated student lists; nested within Learning Management System used by many subjects	Iterative and cyclical (multi-phase) review process; feedback to reviewers	Integration with plagiarism detection, customisable forms; assignment library; flexible assignment of work
Main drawbacks	Calibration and reviews must be on same topic; emphasis on fact-based learning; separate interface and login from LMS	In-house product, no stand-alone version yet	Basic functionality, lack of flexibility; cannot be used at group or pair level; clunky interface	Review criteria minimally customisable and generally restricted to writing (flow, logic and insight); separate site/login; inflexible distribution	Inability to specify distribution rules; does not accommodate group structuring within class; separate site/login; no demo version
Features					
Distribution of work	Random	Customisable rules	Random	Random	Manual, random or student-selected
Review criteria/forms	Customisable; library	Customisable; library	Customisable	Restricted (3 dimensions)	Customisable; library
Feedback to reviewers	No	Yes	No	Yes	No
Review sharing	No	Yes	No	No	Yes
Team member rating	No	Yes	No	No	No
Summary tables	No	Yes	No	No	No
User guide	Yes	Yes	Yes (institution-generated)	Yes	Yes
Ease of use	+	++	-/+	++	+
Availability/cost	Free	Restricted to Unimelb*; free	Institutional purchase of building block	Free	Online; Institutional/departmental licence
Web site	http://cpr.molsci.ucla.edu/	http://disweb.dis.unimelb.edu.au/staff/jonmp/praze	http://www.blackboard.com	http://www.lrdc.pitt.edu/schunn/SWoRD	http://turnitin.com/static/peerreview.html
Reference	Russell 2004	Mulder & Pearce 2007	None	Cho & Schunn 2007	None

*standalone version under development

References and resources

- Biggs, J. and Tang, C. (2007). *Teaching for quality learning at university (3rd edition)*. Berkshire: Open University Press.
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society/ *Studies in Continuing Education*, 22(2), pp.151-168.
- Brew, A. (1999). Towards autonomous assessment: using self-assessment and peer assessment. In S. Brown and A. Glasner (Eds) (1999), *Assessment Matters in Higher education: Choosing and Using Diverse Approaches*. Buckingham: Society for Research into Higher Education & Open University Press, pp. 159-171.
- Brindley, C. and Scofield, S. (1998). Peer assessment in undergraduate programmes. *Teaching in Higher Education*, 3(1), pp.79-87.
- Cheng, W. and Warren, M. (1997). Having second thoughts: student perceptions before and after a peer assessment exercise. *Studies in Higher Education*, 22 (2), pp.223-239.
- Cheng, W. and Warren, M. (2000). Making a difference: using peers to assess individual students' contributions to a group project. *Teaching in Higher Education*, 5(2), pp.243-255.
- Cho, L., Schunn, C.D. and Wilson, R.W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspective
- Cho, K. and Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers and Education*, 48(3), 409-426.
- Dochy, F., Segers, M. and Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: a review. *Assessment and Evaluation in Higher Education*, 24(3), pp.331-351.
- Falchikov, N. (2005). *Improving assessment through student involvement: Practical solutions for aiding learning*. London: RoutledgeFalmer.
- Fallows, S. and Chandramohan, B. (2001). Multiple approaches to assessment: reflections on use of tutor, peer and self-assessment. *Teaching in Higher Education*, 6(2), pp.229-246.
- Gehringer, E.F. (2001). Electronic Peer review and peer grading in Computer-Science courses. *ACM SIGCSE Bulletin*, 22(1), pp.139-141.
- Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown and A. Glasner (Eds) (1999), *Assessment Matters in Higher education: Choosing and Using Diverse Approaches*. Buckingham: Society for Research into Higher Education & Open University Press, pp.41-53.
- Hanrahan, S. and Isaacs, G. (2001). Assessing self- and peer-assessment: the students' views. *Higher Education Research and Development*, 20(1), pp.53-71.
- Liu, N. and Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), pp.279-290.
- Lundstrom, K. and Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second language Writing*, 18, pp.0-43.
- Magin, D. (2001). Reciprocity as a source of bias in multiple peer assessment of group work. *Studies in Higher education*, 26(1), pp.53-64.
- Mulder, R.A. and Pearce, J.M. (2007). PRAZE: innovating teaching through online peer review. In ICT: Providing choices for learners and learning. *Proceedings ASCILITE Singapore 2007: 727-736*.
- Oakley, B., Felder, R. M., Brent, R. and Elhadj, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning*, 2(1), 9-34.
- Pope, N. (2001). An examination of the use of peer rating for formative assessment in the context of theory of consumption values. *Assessment and Evaluation in Higher education*, 26(3), pp.235-247.
- Rubin, R.S. (2006). The academic journal review process as a framework for student developmental peer feedback. *Journal of management Education*, 30(2), 378-398.
- Russell, A.A. (2004). Calibrated Peer Review: a writing and critical thinking instruction tool. In *Invention and impact: building excellence in undergraduate science, technology, engineering and mathematics (STEM) education*. AAAS.
- Sluijsmans, D., Moerkerke, G., Dochy, F., and Van Merriënboer, J. (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation*, 27 (2), pp.153-173.
- Sluijsmans, D.. (2002). Establishing learning effects with integrated peer assessment tasks. Retrieved from <http://www.palatine.ac.uk/files/930.pdf> on 5 July 2009.
- Somervell, H. (1993). Issues in assessment, enterprise and higher education: the case for self-, peer and collaborative assessment. *Assessment and Evaluation in Higher Education*, 18(3), pp.221-233.

- Søndergaard, H. (2009). Learning from and with peers: the different roles of student peer reviewing, *Proceedings of the Fourteenth Annual SIGCSE/SIGCUE Conference on Innovation and Technology in Computer Science Education*, Association for Computing Machinery, pp. 31-35.
- Topping, K. (1998). Peer assessment between students in colleges and universities, *Review of Educational Research*. 68(3), pp.249-276.
- Topping, K., Smith, E.F., Swanson, I. and Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in Higher Education*, 25(2), pp.149-169.
- Van den Berg, I., Admiraal, W. and Pilot, A. (2006). Design principles and outcomes of peer assessment in higher education, *Studies in Higher Education*, 31(3), pp.341-356.

Online Resources

- A discussion of group work, including the assessment of relative individual contribution, with case studies
<http://www.bioassess.edu.au/assessment-types/groupwork>
- A discussion of peer and self assessment, including specific examples
<http://www.bioassess.edu.au/assessment-types/peer-and-self-assessment>
- A short multimedia tutorial on how to give good peer feedback on writing, University of Pittsburgh
<http://www.pitt.edu/~schunn/feedback/>
- Integrating writing into your course: Advice on incorporating peer review in your class, University of Wisconsin
<http://mendota.english.wisc.edu/~WAC/category.jsp?id=20>
- iPeer - Peer Evaluation web application, University of British Columbia Faculty of Applied Science,
<http://ipeer.apsc.ubc.ca/home/>
- Peer editing advice, University of North Texas
http://www.unt.edu/writinglab/peer_editing/index.htm
- Peer review & feedback forms, The University of Hawai'i at Mānoa
<http://www.mwp.hawaii.edu/resources/wm7.htm>
- Teaching guides: using student peer review, Colorado State University
<http://writing.colostate.edu/guides/teaching/peer/>
- The Peer Review Process, Westmont College (Telford Work.net)
<http://www.westmont.edu/~work/material/peerreview.html>
- Turnitin: <http://turnitin.com>