Student peer assessment

Stephen Bostock

Summary

Student assessment of other students' work, both formative and summative, has many potential benefits to learning for the assessor and the assessee. It encourages student autonomy and higher order thinking skills. Its weaknesses can be avoided with anonymity, multiple assessors, and tutor moderation. With large numbers of students the management of peer assessment can be assisted by Internet technology.

Biography

Stephen Bostock FSEDA is Director for IT, Department of Computer Science, at Keele University, where he is also an Academic Staff Developer. After an initial career as a biologist, he developed interests in learning technology and then staff development.

Keywords

Peer assessment, self-assessment, ownership, authentic assessment, reliability, validity, web-based formative assessment

Introduction

Peer assessment is assessment of students by other students, both formative reviews to provide feedback and summative grading. Peer assessment is one form of innovative assessment (Mowl, 1996, McDowell and Mowl, 1996), which aims to improve the quality of learning and empower learners, where traditional forms can by-pass learners' needs. It can include student involvement not only in the final judgements made of student work but also in the prior setting of criteria and the selection of evidence of achievement (Biggs, 1999, Brown, Rust and Gibbs, 1994).

Peer assessment can be considered part of peer tutoring (Donaldson and Topping, 1996). As with other forms of peer tutoring, there can be advantages for both tutor and tutee (Hartley, 1998, 135). Topping (1996, 7) describes the potential advantages of peer tutoring, including the development of the skills of evaluating and justifying, and using discipline knowledge.

Self and peer-assessment are often combined or considered together. They have many potential advantages in common. Peer assessment can help self-assessment. By judging the work of others, students gain insight into their own performance. "Peer and self-assessment help students develop the ability to make judgements, a necessary skill for study and professional life" (Brown, Rust and Gibbs, 1994).

Brown, Rust and Gibbs (1994), Zariski (1996), Race (1998) and others have described some potential advantages of peer assessment for students as

- giving a sense of ownership of the assessment process, improving motivation
- encouraging students to take responsibility for their own learning, developing them as autonomous learners
- treating assessment as part of learning, so that mistakes are opportunities rather than failures

- practising the transferable skills needed for life-long learning, especially evaluation skills
- using external evaluation to provide a model for internal self-assessment of a student's own learning (metacognition), and
- encouraging deep rather than surface learning.

Self and peer assessment "promote lifelong learning, by helping students to evaluate their own and their peers achievements realistically, not just encouraging them always to rely on (tutor) evaluation from on high" (**Brown, 1996**).

Authentic assessment

What is being assessed? The distinctive feature of HE is the learning and assessment of "higher order thinking skills", including assessing and evaluating (Heywood 2000, chapter 2). Ramsden, for example, includes in the aims of higher education independent judgement and critical self awareness (Ramsden, 1992, chapter 3). A common difficulty is that the achievement of such stated aims is undermined by the method of assessment, known as assessment backwash (Biggs, 1999, chapter 8): from the student perspective, assessment defines the actual curriculum (Ramsden, 1992, 187). Any assessment method has wider effects than simple measurement; it can support the achievement of the planned learning outcomes or undermine them. Peer assessment, in this sense, is authentic: making peer assessments involves using discipline knowledge and skills, while accepting peer assessments engages with others' knowledge and skills.

The use of peer assessment encourages students to believe they are part of a community of scholarship. In peer assessment we invite students to take part in a key aspect of higher education: making critical judgements on the work of others. We thus bring together the values and practices of teaching with those of research (Rowland 2000, Boud, 1990).

Feasibility

So far this account has been principled and high-minded. Realistically, though, in the recent U.K. circumstances of worsening staff-student ratios, there is little point is exhorting ourselves to better practice if the increased benefit to students is counter-balanced by an increased need for staff time. But we may avoid this drawback by using the neglected teaching resource of the students themselves. There are possible gains in cost effectiveness; teachers can be managing peer assessment processes rather than assessing large numbers of students directly (Boud, 1995, 42, Race 1998).

We know that learning is improved by detailed, positive and timely feedback on student work (e.g. Brown, Race and Rust 1995, 81). It is therefore worth considering whether student peer-assessment can increase the amounts of feedback that students can derive from their work (**Race, 1995**). An important role for self and peer assessment is providing additional feedback from peers while allowing teachers to assess individual students less, but better. This helps a move from assessing quantity of student work to assessing quality, and the higher order thinking skills (Boud, 1995).

Computer based assessment is another possibility for providing more feedback to students. While giving more predictable feedback, it requires a greater initial investment from staff and is more limited than most students in the quality of its responses.

Possible problems

What are the potential problems of peer assessment? At first sight there must be difficulties with the validity and reliability of assessment done by students. In the case of formative assessment (peer review) we can ask if feedback from fellow students will be accurate and valuable. This will be improved by the use of clear criteria (aligned with the learning objectives, of course), by double anonymity of assessors and assessees, and by having multiple assessors of each piece of work. Ultimately, the value of student assessment will depend on the many variables affecting learning in a specific course.

There is an additional problem with student summative assessment providing grades, and less consensus on its value (Hartley, 1998, 136, **Brown, Rust and Gibbs, 1994, 5.2**). How accurate are peer gradings? In some there is a tendency to under-mark (for example, Penny and Grover, 1996 in Heywood, 2000, 387). Stefani (1994) found student gradings to be closely related to tutor gradings but somewhat lower. Marcoulides and Simkin (1995) found that their students graded accurately and consistently. Boud and Holmes (1995) describe a peer assessment scheme that was "as reliable as normal methods of marking" with a slight bias to over-mark. Haaga (1993) found that students providing double-blind reviews of journal manuscripts were more reliable than professional peer reviews! In general, accuracy is good and is improved where assessment is double-anonymous, assessments are going to be moderated by a tutor, there are clear criteria, and assessors have some experience or training in assessment.

A case

A fairly typical example of formative and summative peer assessment was carried out in 1999/2000 by the author on an MSc module (**Bostock 2000**). 38 students developed instructional web applications (hypermedia or tutorials) on a topic of their choice for 25% of the module assessment. Each student placed their "draft" application on their web space, from which four assessors per assessee provided formative reviews as text criticisms and percentage marks against five criteria. Anonymity of authors was not possible as the student web addresses included their username but the assessors were anonymous; code numbers were used to identify reviews. After receiving anonymous reviews of their work students had time to improve it, and final versions were mounted on the web spaces by a submission deadline. Summative assessments of the same applications were done by the original assessors, sending only marks to tutors, apparently for moderation. The four marks per author were compiled but, in fact, the tutor re-marked all the work.

Sixteen students returned an anonymous evaluation of the assessments. For most students, some or all of the formative reviews had been useful, especially as anonymity allowed some reviews to be "ruthless". Text feedback was valued more than marks. Some said they had wanted more time to act on the criticisms. Most said that seeing other students' work had also been valuable. Feelings were mixed about the use of student summative marking in the module grade, and most only wanted them used if moderated by the tutor. The main problem with the summative assessments was that student preoccupation with the final examination meant that some students did not do them. The marking was variable. Student marks for any one application had a range of 11% with a standard deviation of 6.6%, on average. The correlation between the mean student mark and the tutor mark was only 0.45. This might be improved in future with negotiated criteria (Race, 1998) and more assessment practice (Brown, Sambell and McDowell, 1998).

Managing peer assessment

As with most teaching innovations, peer assessment requires more up-front preparation than the status quo. However, students in the case received more feedback than tutors could have provided, and found it useful. The coursework was of a higher standard than previous years (see also Sims, 1989). No tutor time was saved on summative marking because student marks were not used, being too unreliable.

The main practical difficulty was generating the messages of instructions to students and assessments between students, and archiving them for possible auditing. Technology did help. All students had web spaces into which they placed their work, for viewing by assessors. Email was used for all messages and web forms were used to collect and send anonymous assessments. The allocation of assessors was done randomly but with an equal number of assessments per student and avoiding assessor-assessee pairs. Monitoring the progress of assessments meant checking the message archive. Many of these processes could be further automated.

There are few examples of computer support for managing peer assessment. MacLeod (1999) had students use generic communication software during formative reviewing. Wood (1998) discussed electronic peer refereeing of journal articles. Robinson (1999) set out the requirements of a system to support student peer assessment, and one is being developed at Keele to support the MSc module, with the intention of making it publicly available later.

References

Biggs, J. 1999 *Teaching for Quality Learning at University*, Buckingham: SRHE and Open University Press

Bostock, S.J. 2000 *Computer Assisted Assessment - experiments in three courses*. A workshop at Keele University, May.
http://www.keele.ac.uk/depts/cs/Stephen_Bostock/docs/caa-ktn.htm

Boud, D, 1990 Assessment and the promotion of academic values, *Studies in Higher Education*, 15(1), 101-111

Boud, D. 1995 Assessment and learning: contradictory or complimentary? 35-48 in Knight, P. (ed.) *Assessment for Learning in Higher Education*, London: Kogan Page/SEDA

Boud, D. and Holmes, H. 1995 Self and peer marking in a large technical subject, 63-78 in Boud, D. *Enhancing Learning through Self Assessment*, London: Kogan Page

Brown, S. 1996 *Assessment*, in DeLiberations
http://www.lgu.ac.uk/deliberations/assessment/invite.html

Brown, S., Race, P. and Rust, C. 1995 Using and experiencing assessment, 75-86 in Knight, P. (ed.) *Assessment for Learning in Higher Education*, London: Kogan Page/SEDA

Brown, S., Rust, C. and Gibbs, G. 1994 Involving students in the assessment process, in *Strategies for Diversifying Assessments in Higher Education*, Oxford: Oxford Centre for Staff Development, and at *DeLiberations*
http://www.lgu.ac.uk/deliberations/ocsd-pubs/div-ass5.html

Brown, S., Sambell, K. and and McDowell, L. 1998 What do students think about assessment? 107-112 in *Peer Assessment in Practice*, Brown, S. (ed.) (SEDA paper 102) Birmingham: SEDA

Donaldson, A.J.M. and Topping, K.J. 1996 *Promoting Peer Assisted Learning amongst Students in Higher and Further Education* (SEDA paper 96) Birmingham: SEDA

Haaga, D.A.F. 1993 Peer review of term papers in graduate psychology courses, *Teaching of Psychology*, 20 (1), 28-32

Hartley, J. 1998, *Learning and studying*, London: Routledge

Heywood, J. 2000 *Assessment in Higher Education*, London: Jessica Kingsley Publishers

MacLeod, L. 1999 Computer aided peer review of writing, *Business Communication Quarterly* 62 (3) 87-94

Marcoulides, G.A. and Simkin, M.G. 1995 The consistency of peer review in student writing projects, *Journal of Education for Business* 70 (4) 220-223

McDowell, L. and Mowl, G. 1996 Innovative assessment - its impact on students, 131-147 in Gibbs, G. (ed.) *Improving student learning through assessment and evaluation*, Oxford: The Oxford Centre for Staff Development

Mowl, G. 1996 Innovative Assessment, in DeLiberations
http://www.lgu.ac.uk/deliberations/assessment/mowl_content.html

Race, P. 1995 The Art of Assessing, *New Academic*, Autumn 1995, 3-5 and Spring 1996, 3-6 and in DeLiberations
http://www.lgu.ac.uk/deliberations/assessment/artof_fr.html

Race, P. 1998 Practical Pointers in Peer Assessment, 113-122 in *Peer Assessment in Practice*, Brown, S. (ed.) (SEDA paper 102) Birmingham: SEDA

Ramsden, P. 1992 *Learning to teach in Higher Education*, London: Routledge

Robinson, J.M. 1999 Computer-assisted peer review, in Brown, S., Race, P and Bull, J (eds.) *Computer-assisted Assessment in Higher Education*, London: Kogan Page/ SEDA

Rowland, S. 2000 *The Enquiring University Teacher*, Buckingham: SRHE and Open University Press

Sims, G. K. 1989 Student peer review in the classroom: a teaching and grading tool, *J. Agron. Educ.* 8 (2) 105-108.

Stefani, L.A.J. 1994 Peer, Self and tutor assessment: relative reliabilities, *Studies in Higher Education*, 19 (1) 69-75

Topping, K. 1996 *Effective Peer Tutoring in Further and Higher Education*, (SEDA Paper 95) Birmingham: SEDA

Wood, D.J. 1998 Peer review and the web: the implications of electronic peer review for biomedical authors, referees and learned society publishers, *Journal of Documentation* 54 (2) 173-197

Zariski, A. 1996 Student peer assessment in tertiary education: Promise, perils and practice. In Abbott, J. and Willcoxson, L. (Eds), *Teaching and Learning Within and Across Disciplines*, p189-200. Proceedings of the 5th Annual Teaching and Learning Forum, Murdoch University, February 1996. Perth: Murdoch University.
http://cleo.murdoch.edu.au/asu/pubs/tlf/tlf96/zaris189.html

Acknowledgements