Investigating the potential of self and peer assessment to develop learning oriented assessment tasks

Keith Willey
University of Technology, Sydney, Australia
Keith.Willey@uts.edu.au

Anne Gardner
University of Technology, Sydney, Australia
Anne.Gardner@uts.edu.au

Abstract: *The authors have previously reported the effectiveness of using self and peer assessment to improve learning outcomes in groupwork by providing opportunities to practise, assess and provide feedback on students' attribute development. Combining this research and that reported in the literature on learning-oriented assessment we theorised that self and peer assessment would be an ideal tool to develop and facilitate assessments specifically designed to promote learning.*

In this paper we report testing this theory by integrating self and peer assessment into different learning oriented assessment tasks within a single subject. These tasks use self and peer assessment to not only assess a student's contribution to a team project but also assess individual student assignments, their understanding and judgement. The results demonstrate that self and peer assessment is an extremely effective tool in facilitating learning orientated assessments.

Introduction

In addition to providing fairer assessment of group work, self and peer assessment is reported as assisting students to develop important professional skills including reflection and critical thinking (Mello, 1993; Somervell, 1993). Michaelsen discusses the use of self and peer assessment to promote peer learning (Michaelsen et al., 2004), while Willey and Freeman (2006a, 2006b) report using it to produce formative learning-oriented feedback to complete the learning cycle and encourage the ongoing development of skills. The positive effects of peer and self assessment on student learning are also reported by Hanrahan and Isaacs (2001). Furthermore Boud and Falchikov (2007) discuss its use for developing students' skills for lifelong learning. More recently the authors have reported the effectiveness of using self and peer assessment to improve learning outcomes by providing opportunities to practise, assess and provide feedback on students' attribute development (Willey & Gardner, 2008a).

In the last decade various researchers have suggested that assessment had to change from 'assessment of learning' to 'assessment for learning' (Torrance 2007). It was recommended that assessment practices be developed to support learning and build, not undermine, student confidence, achievement and progress (Black & Wiliam, 1998, Gibbs & Simpson, 2004, Brown, 2004). Learning-oriented assessment embeds learning in assessment, reconfiguring its design to emphasise the function of learning (Keppell & Carless, 2006; Keppell et al, 2006). It has three main elements, assessment tasks that also focus on learning, involving students in the assessment process to develop their attributes including judgement and feed-forward to improve subsequent contributions and learning (Carless, 2007, Black & Wiliam, 1998).

However, one must be careful in designing assessments, as good intentions can lead to bad practice. While the provision of detailed feedback and assistance by instructors increases the likelihood that students will succeed, care needs to be taken that the challenge for the student is not removed from the learning process and/or that the quality and validity of the assessment outcome is not reduced (Torrance 2007). The authors have noticed a tendency for students to become "incremental learners" whereby they seek ongoing feedback and advice from their instructors to improve their submission. In these instances there is a danger that only the instructors are exercising judgement, with students simply implementing without reflection or engagement what the instructor has told them to do, with the narrow focus of securing a better grade. Rather than building the skills required for independent learning such practices may actually produce "…students who are *more* dependent on their tutors and assessors (Torrance 2007, p. 282). This is in contrast to peer learning which encourages students to take responsibility for their own learning (Keppell et al, 2006).

In response to this research we theorised that self and peer assessment would be an ideal tool to develop learning oriented assessments that would also address the above issues. In particular we were interested in making students more responsible for their own learning by requiring them to provide their own feedback and contribute to their own self assessment and to the assessment of their peers.

In this paper we report testing this theory by integrating self and peer assessment into a number of different specifically designed learning oriented assessment tasks within a single subject. Self and peer assessment facilitated using the online tool SPARK^{PLUS} (Willey & Gardner, 2008c) was used to assess and provide feedback on individual student submissions, student's judgement through benchmarking exercises and to assess and provide feedback on contributions to a team project.

SPARK^{PLUS}

SPARK^{PLUS} is a tool for facilitating the use of self and peer assessment. It has the capacity to not only assess a student's contributions to a team project, but also allows students to self and peer assess individual work and improve their judgment through benchmarking exercises (Willey & Gardner, 2008a; Willey & Gardner, 2008c).

SPARK^{PLUS} assists students to make their self and peer assessments by requiring them to rate each other over multiple criteria (Figure 1). The program has the capacity to produce three assessment factors:

- The Self and Peer Assessment (SPA) factor is a weighting factor determined by both the self and peer rating of a student's contribution. It is typically used to change a team mark for an assessment task into an individual mark as shown below:

  Individual mark = team mark * Individual's SPA

- The Self Assessment to Peer Assessment (SAPA) factor. This is the ratio of a student's own rating of themselves compared to the average rating of their contribution by their peers. The SAPA factor has strong feedback value for development of critical reflection and evaluation skills eg. a SAPA factor greater than 1 means that a student has rated their own performance higher than the average rating they receive from their peers and vice versa.
- The third factor is a percentage mark, the calculation of which depends on the type of task that has been selected (e.g. benchmarking exercise or marking individual work).

SPARK^{PLUS} also allows students to provide anonymous written feedback to their peers and provides a number of options for graphically reporting results.

Design Fundamentals

Design Fundamentals is a Stage 3 compulsory core subject within all Engineering Degrees at the University of Technology, Sydney. The subject's typical cohort is approximately 300 students with tutorial classes being limited to a maximum of 32 students.

The subject's primary aims are to:

1. Develop students' understanding of the engineering design process
2. Provide students with the skills to develop a small engineering project from initial concept to the production of a prototype.
3. Continue the development of students' professional skills including teamwork, critical evaluation, feedback and communication commenced in earlier subjects.

To promote the development of professional skills, provide students with feedback, improve students' judgement and critical evaluation skills and encourage academic honesty, a process of self and peer assessment (collected using the online tool SPARK^{PLUS}) is used four times during the semester.

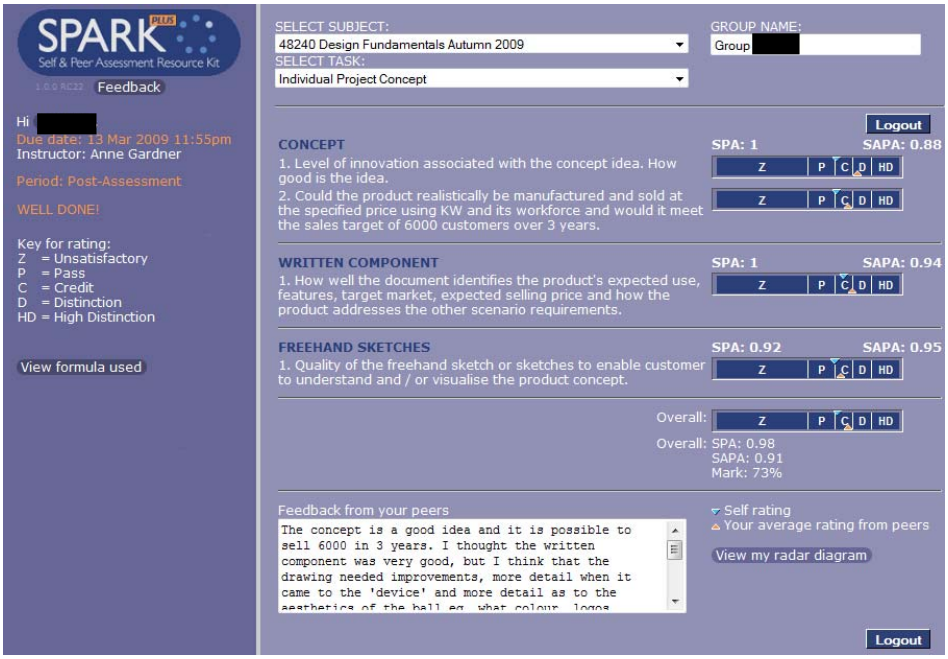

Figure 1: A student's SPARK^{PLUS} results screen for a task where each student had to self assess their own submission and peer assess the individual submissions of their team peers.

Method

In line with Carless (2007) we designed assessments tasks using self and peer assessment to:

- Develop assessment tasks to be learning tasks
- Involve students in the assessment process to develop their professional skills including judgement
- As much as possible allow feedback to be fed forward to improve subsequent assessment and learning

Self and Peer assessment was integrated into four distinct collaborative learning assessment tasks that, when combined, form a major design project. The tasks were as follows:

Individual Project Concept: Students use SPARK^{PLUS} to assess their own and seven of their peers' individual project concept against a number of specified criteria. In the next tutorial the group of eight students debate the merits of each individual submission (discussing their individual strengths and weaknesses) and collectively place them in order from best to worst awarding a mark for each one. Students then receive the results from SPARK^{PLUS} and are asked to reflect on any differences between the results produced from their individual assessments (SPARK^{PLUS}) and those produced collectively in their peer group. The tutor marks the best report from each group (as identified by the students) and determines marks for the other reports using the weighting produced by SPARK^{PLUS}.

The peer learning groups are divided into two groups of four students. Each group works together to complete the three remaining stages of the project.

Benchmarking Exercise: Students are provided with a Sample Requirement Specification report. After discussing the marking criteria each student has to grade the report using SPARK^{PLUS}. In their next tutorial each group of four students discuss their individual marking of the report and re-mark it collectively against the criteria. Students then re-combine into their peer learning groups (two groups of four students) and discuss their group's marking of the report, reflecting on any differences and collectively re-mark it. Tutors then discuss how the academic marked the report. After the tutorial students may log on to SPARK^{PLUS} and compare their individual marking to the academic's marking of the report for each individual criterion and read the academic's comments. In addition, SPARK^{PLUS} produces a weighted mark related to how close the student's individual assessment was to the academic's assessment.

Both of these tasks start with students working individually. This individual work is assessable through SPARK^{PLUS}. Having individual assessable work as the first part of this process promotes students coming prepared to the following collaborative activities. Even though the collaborative learning phase is not directly assessable there is incentive for the students to participate. The individual project concept activity helps students to determine the 'best' idea for the group to pursue as their semester project, while the benchmarking exercise assists students to write a 'better' requirements specification report.

Requirement Specification: each group of students produces a requirement specification for their design project. Students use SPARK^{PLUS} to rate their own and their team peers' contribution to this stage of the project. The SPARK^{PLUS} SPA factors are used to produce individual marks by moderating the mark for the group's submission. In the next tutorial the group's individual results are distributed to all group members and discussed. Groups are guided through a feedback process. This process begins with self evaluation where students share with their group what they have learnt or discovered about their strengths, weaknesses or performance from the exercise. Students are encouraged to identify how they could improve their own performance and in what way they would approach the task differently if they had to do it again. Students are asked to suggest how others in their group may have approached their tasks differently to achieve a better group result, how aspects of their behaviour affected the team and the benefits of changing that behaviour, and to reflect on how team peers could have learnt more from the process. The in-class discussion concludes by teams agreeing how to improve their overall team and individual performance for the remaining parts of the project and /or in future group work opportunities.

Project Report, Oral Presentation and Prototype Demonstration: each group of students produce a project report, make an oral presentation and present their prototype design. Students again use SPARK^{PLUS} to rate their own and their team peers' contributions to this stage of the project. This is followed by the same feedback process and discussion as previously described.

In the last semester of 2008 three subject surveys were conducted to assess the effectiveness of the self and peer assessment processes used in the subject to facilitate learning orientated assessment tasks, encourage students to actively engage in their own learning and facilitate high level learning outcomes. The questions were a mixture of free response and 4 point Likert format. All students undertaking the project (eligible cohort 255) were required to participate in the self and peer assessment exercises. The first two surveys (Individual Project Concept and Benchmarking) were conducted in tutorial classes resulting in 209 and 201 students responding respectively. In the benchmarking survey one student only completed the background questions and hence their submission was excluded resulting in a participating cohort of 200. The post subject survey was much longer (60 questions), conducted online, and was completed by 89 students from an eligible cohort of 255 (35%).

Results

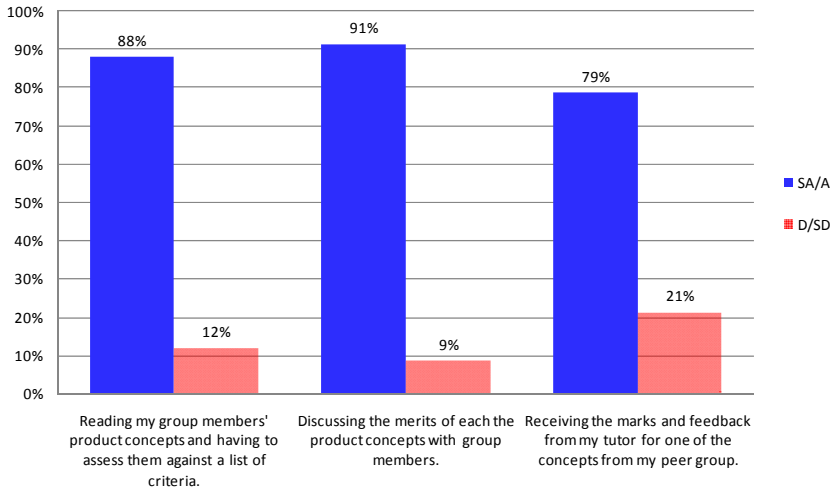

Figure 2: Student survey results for Self and Peer Assessment Marking of Individual Project Concepts in response to the question "*My ability to choose a product concept and write a concept document to meet a list of requirements increased as a result of:*"

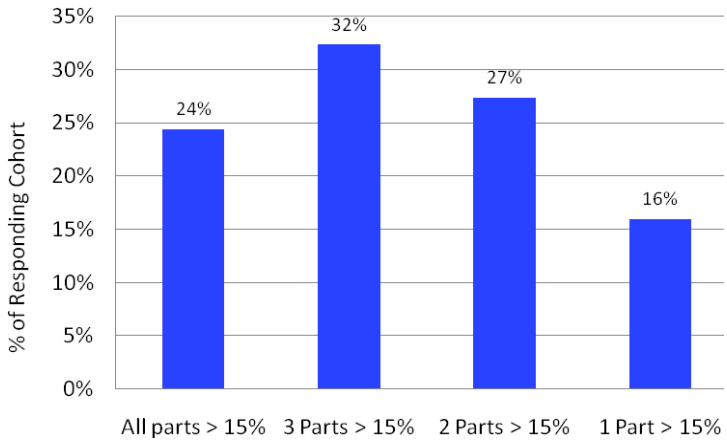

Figure 3: **Student survey results for Self and Peer Assessment Marking of Individual Project Concepts in response to the question: *If you consider the amount that your understanding / ability to write a quality requirement specification increased 100% as a result of this exercise, how much did each of the following contribute to improving your understanding / ability?***

a) ***Having to read others' reports by myself & assessing them against the criteria (the assessments you entered via SPARK).***
b) ***Discussing the different concepts in the group.***
c) ***The feedback were received from our tutor as they explain the marking of an exemplar concept.***
d) ***The feedback I received about my concept from the SPARK^{Plus} group radar diagrams and factors.***

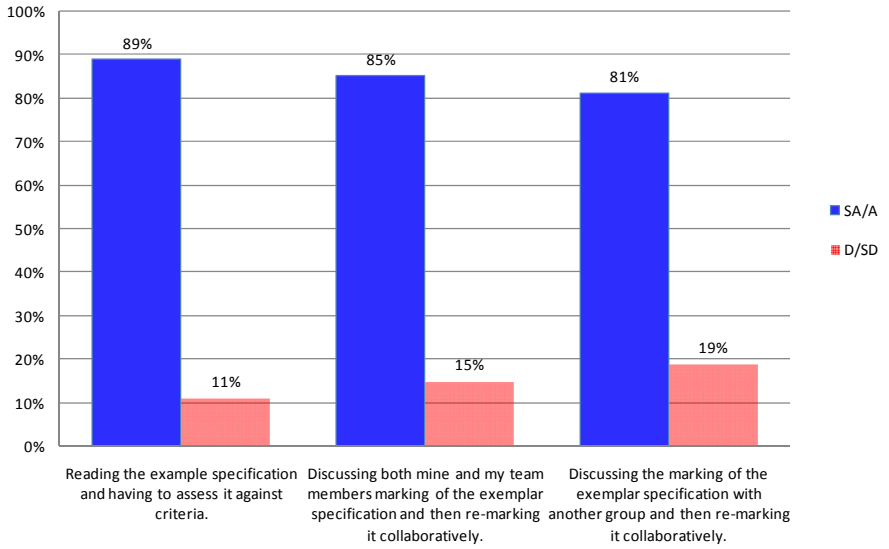

Figure 4: Student survey results for Self and Peer Assessment Benchmarking Exercise in response to the question "My ability to write a quality requirement specification has increased as a result of:"

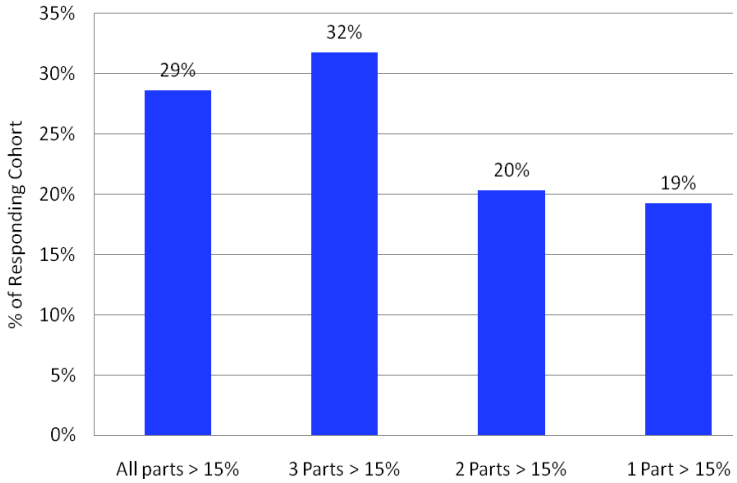

Figure 5: **Student survey results for Self and Peer Assessment Benchmarking Exercise in response to the question *If you consider the amount that your understanding / ability to write a quality requirement specification increased 100% as a result of this exercise, how much did each of the following contribute to improving your understanding / ability?***

a) ***Having to read the specification by myself & assessing it against the criteria (the assessments you entered via SPARK)***
b) ***Discussing the specification marking within my group and then marking it collaboratively***
c) ***Discussing the specification marking within the combined group and then marking it collaboratively***
d) ***The feedback, guidance and explanation the class received from our tutor***

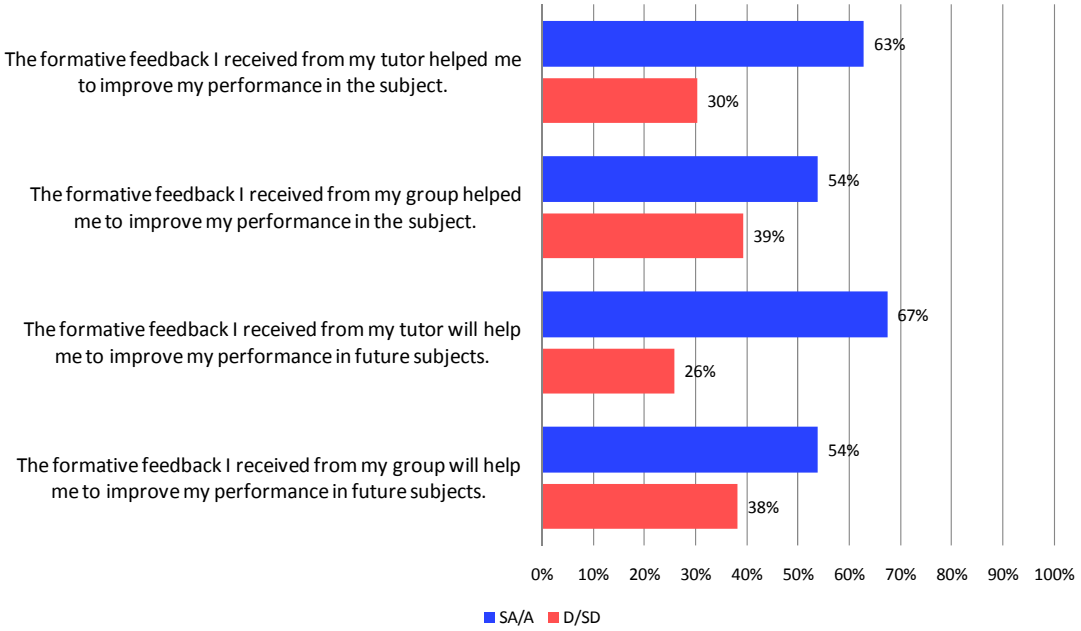

Figure 6: Results from Post Subject Survey

Discussion

Figures 2 - 6 present the survey results relevant to this paper. The 'Strongly Agree' and 'Agree' responses were combined to give an aggregate result, as were the 'Strongly Disagree' and 'Disagree' responses. The percentage of any unanswered questions are generally not shown but can be calculated by subtracting the provided results from 100%.

The results (Figure 2 and 4) show that the majority of students (ranging from 79% to 91%) felt that all aspects of the group marking of individual submissions and the benchmarking exercises improved their ability to meet the prescribed learning outcomes. While there were some complaints from students that it took too long to complete all the parts of these exercises, generally speaking most students were positive in line with the survey free response comments below:

Peer review: *"Allows you to see what people think of your work and how you can improve"* [sic].
Benchmarking: *"Reviewing and marking a previous piece of work helped to understand the theory from the lectures. Knowing we need to write a Requirements Specification that is unambiguous is easy enough to know, but WHAT that actually looks like, and doing it is hard. Getting a picture of what NOT to do first, helps developing that knowledge"* [sic].

Furthermore the results in figure 3 and 5 show that 56% (group marking of individual submissions) and 61% (benchmarking) of responding students reported that three of the four tasks within these exercises contributed more than 15% to their improved understanding and/or ability to meet the relevant learning outcomes. This is a positive result especially given that only one of the four integrated activities was actually assessable and that the assessable part of the task was organised to run outside of scheduled class times. The remaining tasks within the learning orientated assessments were conducted within normal tutorial sessions. Thus more than 50% of responding students reported that at least 30% of their improved understanding and/or ability to meet the prescribed learning outcomes was a result of the non-assessable components of the learning oriented tasks. This result is in line with Gibbs and Simpson (2004) "…coursework does not have to be marked to generate the necessary learning"(p.8).

The results also indicate that the use of self and peer assessment made a significant contribution to students' learning and their ongoing development for the exercises used to determine their contribution to the last two stages of the project. The results presented in Figure 6 demonstrate that we were successful in designing assessment tasks that also promoted learning. For example the

feedback processes built into the assessment tasks were not assessed. While students who did not attend the feedback sessions received a small deduction from their project mark, the quality of the feedback that they provided through their self and peer assessments and their engagement in the feedback process were not directly assessed. Despite this students engaged with the learning opportunity provided, with 54% agreeing that the formative feedback they received from the group would not only help them to improve their performance in Design Fundamentals but also in future subjects. This is a strong result when compared to the benefits reported by students in regard to the feedback they received from their tutors. Tutor feedback was regarded by more respondents as being useful to their ongoing learning with 63% reporting tutor feedback would help them to improve their performance in Design Fundamentals and 67% agreeing that it would help them in future subjects.

As our survey results do not ascertain how useful the feedback was, only that it **was** useful we cannot make a direct comparison as to how effective student feedback was in comparison to that received from their tutors. However, the significance of the fact that just over 50% of students reported that activities within the assessment task contributed to their ongoing learning cannot be underestimated. This is feedback and learning provided by students which would not have been provided if the assessment task was not designed to be learning oriented. Furthermore, using an online tool to collect, distribute and report this feedback means that these learning gains could be made even in a large class without undue academic effort. It is our belief that by deliberately designing the assessment aspect of the tasks to be conducted outside of schedule classes, and organising the learning orientated component to be conducted within class time significantly increased student engagement and the benefits they received.

In addition to the reported results there were many positive free response comments typified by those provided below:

"Peer assessment facilitated by SPARK improved my group work experience by facilitating and giving me peer feedback with regards to the contributions by the team. It gave all team members an opportunity to give fair and constructive feedback (mostly) to each other, thus improving the performance in projects throughout the semester, and most likely in later subjects also."

"Improved my group work experience as SPARK enables a fairer assessment, I was driven to participate and function with my team as a group. It gave me the opportunity to see my effort (by my SPA rating) and also to know what other team members thought about my performance from feedback received. I really enjoyed working in a group for this subject and I think SPARK had a big influence in that" [sic].

While the results clearly demonstrate the potential of self and peer assessment as an effective tool to facilitate learning oriented assessments we cannot escape the fact that '*from our students point of view, assessment always defines the actual curriculum'* (Ramsden, 2003, p. 182, emphasis added). This view is supported by the following free response comment provided by a student in relation to the tutorial feedback sessions within the subject:

"Feedback couldn't be used to improve mistakes and consequently improve the assessment marks. I feel its a big waste when this is the case as the feedback isn't taken as serious as it should be as you cant use it to improve your marks. Even though it helps you to learn, as it doesn't show through in the assessment marks which is ultimately the students number 1 aim, …."(sic).

We are currently in the process of working with students to redesign our assessment tasks to encourage further learning and engagement with the feedback provided. In particular we are considering providing students with an opportunity to use the peer feedback they receive to improve their submissions before they are graded. Currently we favour a form of peer review. While regular feedback provided by academic instructors has a danger of encouraging students to become incremental learners, peer review requires students to use their own judgement in determining both the quality of the feedback they receive and how they should respond. Furthermore, we believe that while students may be content to submit substandard work to their instructors, potential embarrassment will motivate them to improve the standard of work they submit to their peers. It is hoped this motivation will encourage students to embrace the challenge of learning, not just focus on learning what is assessed.

Conclusion

The results show that self and peer assessment is an effective tool in facilitating learning orientated assessments. The majority of respondents, greater than 69%, reported that its use improved their ability to meet the required learning outcomes. We are currently in the process of working with students to redesign our assessment tasks to encourage further learning and engagement with the feedback provided. Currently we favour a form of peer review. While regular feedback provided by academic instructors has a danger of encouraging students to become incremental learners, peer review requires students to rely on their own judgement in determining both the quality of the feedback they receive and how they should respond.

References

Black P. & Wiliam D. (1998) Assessment and classroom learning *Assessment in Education: Principles, Policy & Practice; Mar1998, Vol. 5 Issue 1,* p.7

Boud D., and Falchikov, N.(2007) *Rethinking Assessment in Higher Education Learning for the Longer Term.* Routledge.

Brown S. (2004) Assessment for Learning *Learning and Teaching in Higher Education, Issue 1, 2004-05*. pp. 81 – 89.

Carless D. (2007) Learning-oriented assessment: conceptual bases and practical implications *Innovations in Education and Teaching International Vol. 44, No. 1, February 2007,* pp. 57–66

Falchikov, N., and Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70(3), 287-322.

Freeman M. and McKenzie J. (2002), SPARK, A Confidential Web-Based Template for Self and Peer Assessment of Student Teamwork: Benefits of Evaluating across Different Subjects, *British Journal of Educational Technology*, vol. 33, pp. 551-69.

Gibbs G. & Simpson C. (2004) Conditions Under Which Assessment Supports Students' Learning *Learning and Teaching in Higher Education, Issue 1, 2004-05. pp. 3 – 31.*

Hanrahan S. & Isaacs G. (2001) Assessing Self- and Peer-assessment: the students' views *Higher Education Research & Development, Vol. 20, No. 1* pp. 53 -70.

Keppell M. & Carless D. (2006) Learning-oriented assessment:a technology-based case study *Assessment in Education Vol. 13, No. 2, July 2006*, pp. 179–191

Keppell. M., Au E., Ma A., & Chan C., (2006) Peer learning and learning-oriented assessment in technology-enhanced environments *Assessment & Evaluation in Higher Education Vol. 31, No. 4, August 2006,* pp. 453–464

Mello J. A. (1993). Improving individual member accountability in small work group settings, *Journal of Management Education*, vol. 17(2), pp. 253-259, 1993.

Michaelsen L, Knight A., Fink L., (2004) Team-based Learning – A transformative use of small groups in college teaching. USA, Sylus Publishing.

Ramsden P., (2003) *Learning to Teach in Higher Education*, 2nd ed. London: Routledge,

Somervell H. (1993). Issues in assessment, enterprise and higher education: the case for self-, peer and collaborative assessment, *Assessment & Evaluation in Higher Education*, vol. 18, pp. 221–233, 1993.

Torrance H. (2007) Assessment *as* learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education Vol. 14, No. 3, November,* pp. 281–294

Willey, K. and Freeman M. (2006a), Completing the learning cycle: The role of formative feedback when using self and peer assessment to improve teamwork and engagement. *Proceedings of the 17th Annual Conference of the Australasian Association for Engineering Education*, 10 -13th December 2006, Auckland, New Zealand.

Willey K, and Freeman M. (2006b), "Improving teamwork and engagement: the case for self and peer assessment", *Australasian Journal of Engineering Education*. Online publication 2006-02 http://www.aaee.com.au/journal/2006/willey0106.pdf

Willey, K and Gardner, A. (2008a) Using Self Assessment to Integrate Graduate Attribute Development with Discipline Content Delivery. *Proceedings of the 36th Annual Conference of the European Association of Engineering Education (SEFI)* 2-5 July, Aalborg, Denmark.

Willey K., Gardner A., (2008b) Using self and peer assessment for professional and team skill development: do well functioning teams experience the benefits? *Proceedings of the ATN Assessment Conference – Engaging Students in Assessment,* November, 2008. South Australia.

Willey K., Gardner A., (2008c) Improvements in the self and peer assessment tool SPARK: Do they improve learning outcomes? *Proceedings of the ATN Assessment Conference – Engaging Students in Assessment, November, 2008. South Australia.*

Acknowledgements

The initial redevelopment of SPARK was a joint research project between the University of Technology, Sydney and the University of Sydney. The main developers were Dr Keith Willey UTS, A/Prof Mark Freeman USyd (also chief architect and developer of the original SPARK) and Mr Darrall Thompson UTS. We would also like to acknowledge the contribution of Mr Mike Howard who has continued to work with the authors in developing SPARK^{PLUS}.