The Effectiveness of Peer Assessment in Student Learning

Eamonn Ambrose (eamonn.ambrose@ucd.ie)
Michael Smurfit Graduate Business School/
University College Dublin

Abstract

This study reviews the effectiveness of a peer assessment exercise in a master's level module in Supply Chain Operations. The exercise was based on both individual and group peer assessment of an existing formative assignment within the module as well as parallel summative assessment from the lecturer. Students benefited from carrying out the assessment of their peers' work, and from giving and receiving feedback on the work. However, counter to much of the teaching and learning literature, results suggest that students considered the lecturer assessment feedback to be the most valuable learning element of the process.

Keywords: Peer Assessment, Teaching and Learning, Formative Assessment

Introduction

Assessment of learning outcomes is a challenging and resource intensive aspect of higher education teaching and is coming under increased scrutiny in terms of its value as a learning mechanism, its effectiveness in measuring and certifying achieved learning outcomes, and its efficiency in terms of resource requirements These challenges are particularly relevant in the field of operations and supply chain management where assessment methods used are still predominantly traditional and exam oriented (Ambrose, 2016).

In both bachelor's and master's level programmes students are required to achieve higher level learning outcomes including analytical, critical and creative skills as well as metacognitive knowledge. The teaching and certification of these skills requires a range of assessment techniques beyond the traditional exam focused approach. Teachers need to be conversant with the range of assessment methods available and trained in the design and implementation of these methods in a range of educational settings. This study aims to gain insights into the design and implementation of peer assessment as well as its effectiveness as a learning process/mechanism.

Assessment and Learning

Assessment is typically categorised as summative (with a view to measuring and certifying student performance) or formative (with a view to enhancing student learning), often referred to as assessment of learning and assessment for learning respectively (National Forum, 2016). The National Forum also identified a third category – assessment as learning – where the assessment experience is designed to enhance the

students' skills as lifelong independent learners. However, assessment is not a simple dichotomy or trichotomy as often a single assessment exercise can comprise of two or even all three dimensions concurrently (Hernandaz, 2012). Hence it is a challenge to understand how a particular assessment contributes to learning across these multiple dimensions.

Assessment can be designed to meet short-term learning outcomes such as understanding of a particular theory application of an analytical method. However, growing area of interest is learning-oriented assessment (LOA), which considers assessment as a contributor to lifelong learning or self-regulated learning (Hernandez, 2012). In LOA students are active in the assessment process rather than passive receivers of feedback. Assessment can be designed to enhance the students' evaluative skills, and to promote engagement and action targeted at lifelong learning (Carless, 2007). To achieve LOA, Carless (2007) identifies three principles: assessment tasks should be designed to encourage appropriate learning practices appropriate to the context; the students should be engaged with the quality of their own or their peers' work and with the performance criteria for the activity; feedback should be prompt and oriented towards lifelong learning.

It is useful to consider the psychological process whereby external feedback leads to self-regulated learning (Figure 1). For any given task set by the lecturer, students will interpret the challenge and set their own goals, which may or may not align with the lecturer's intended goals. Effective feedback provided on the outcome of the task, e.g. the report or presentation, will enable students to improve their domain and strategy knowledge, thus building both module-specific and lifelong learning skills (Nicol and Macfarlane-Dick, 2006). Effective feedback also has been shown to have a direct impact on student motivation, in particular intrinsic motivation (Tseng and Tsai, 2010).

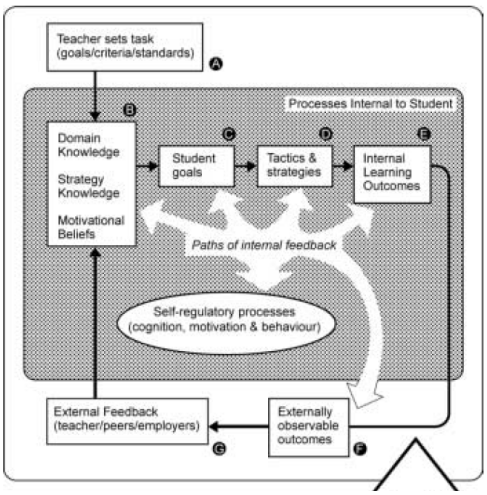

Figure 1: A Model of Self-Regulated Learning, from Nicol and Macfarlane-Dick, 2006

For an effective feedback process, the student must be able to interpret the feedback in a meaningful way and internalise it before it can have an impact of their future learning. This raises a challenge for the lecturer as to how best to provide feedback that will enable students to learn. As the subject matter expert (and as the arbiter of performance standards), the lecturer is well placed to identify gaps between desired and actual performance. However, subject matter experts tend to underestimate the potential performance of subject matter novices, and lecturers don't necessarily have the skills required to communicate the nature of these performance gaps and identifying solutions

(Cho and MacArthur, 2010). There is evidence of wide discrepancies between how lecturers view feedback given, and how students make use of the feedback received. Students generally find that feedback was of limited value in enhancing learning, and often do not even read the feedback given focussing instead on any grades awarded (Gibbs and Simpson, 2004).

It is generally agreed that well-designed assessment activities can help students identify themselves as learners and develop metacognitive skills such as self-awareness, judgement and reflexivity (Boud and Falchikov, 2007). However, there is still debate about how this learning takes place, and hence how to design assessment exercises in order to foster the desired learning outcomes.

Peer Assessment

Peer assessment is defined as a set of activities through which individuals make judgements about the work of others. It is an umbrella term which can include the analysis of the work of others, the judgement of that work in terms of content or grade, and the provision of feedback (commentary and/or grade) to the producer of the work. Effective peer assessment will contribute to student learning in three ways: by enhancing students' goal awareness, as they gain an understanding of exactly what the goals are and an appreciation of different levels of quality of outputs; by improving self-awareness through providing students with the skills and framework with which to assess their own work; and by supporting gap closure activity through students both giving and receiving advice on how to improve their work (Rienholz, 2015). It is important to create a clear distinction between the learning aspects of the peer assessment and any summative assessment activity, so that student can feel free to be critical without it having a negative effect on their peers' grades. It is essential that the peer assessment activity be framed clearly in the context of long-term learning (Boud and Falchikov, 2006). This has led some to argue that formative assessment should be limited to peer feedback and not include peer grading. However, careful integration of feedback and grading can result in greater learning outcomes for students (Liu and Carless, 2006). The student learning takes place progressively over the various stages of the peer assessment process – task engagement, peer analysis, feedback provision and feedback reception. This process can continue through to peer conferencing and revision if appropriate (Reinholz, 2015).

Studies have shown that novices in a subject area find peer assessment comments easier to incorporate into future work and learning than subject matter expert comments. Furthermore, feedback from multiple peers can facilitate even more extensive learning and performance improvement (Cho and MacArthur, 2010). In addition to peers providing feedback commentary on a piece of work, learning from peer grading has been enanced where the lecturer provides a grading rubric for students to apply (Jonsson and Svingby, 2007). Peer grading has been found to correlate quite well with lecturer grading where rubrics are well designed, and the students are taught how to apply them (Liu and Carless, 2006).

By actively involving students in the assessment process, peer assessment can motivate students and extend their learning beyond the confines of the subject matter to encompass lifelong learning skills such as reflection, analysis and diplomatic criticism (Falchikov, 1995). This review leads us to develop a theoretical framework for the study of the learning process steps and learning outcomes associates with peer assessment (Figure 2).

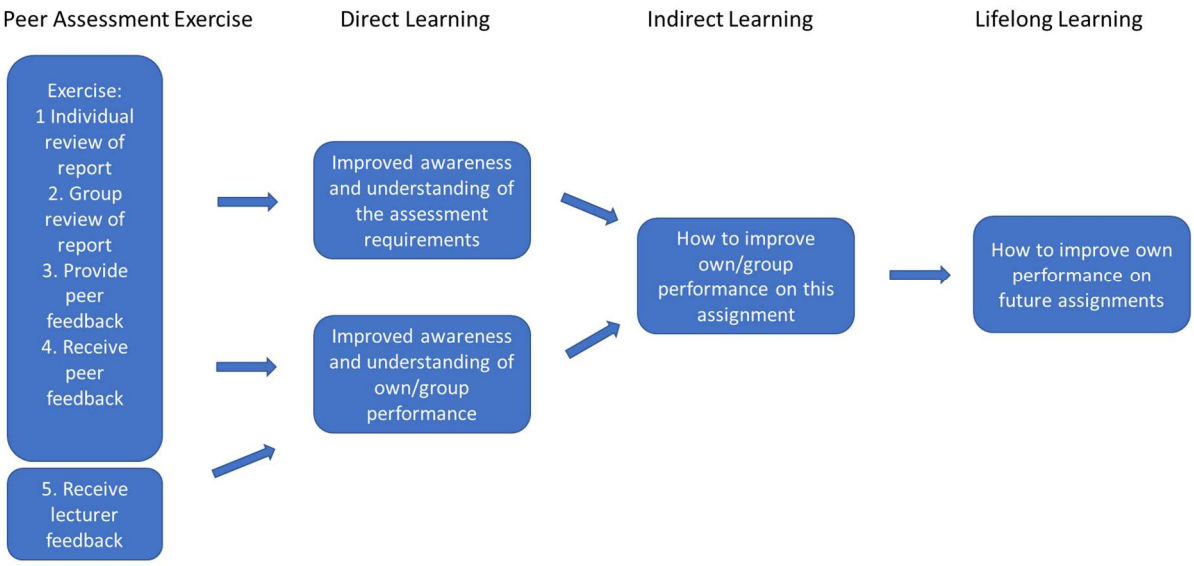

Figure 2: A theoretical framework – process steps and outcomes

Research Questions

We adopt the process approach to peer assessment and consider how each process step contributes to the student learning. At the same time student learning is a multidimensional concept spanning both topic specific content and lifelong learning skills. Hence, we ask

1. How does each of the five process steps contribute to each of the four potential learning outcomes?

Peer assessment is also considered to be as effective as and often better than teacher assessment in that the feedback given is more relevant to students and more easily understood and applied. Hence, we ask

2. How does the peer grading (both individual and group) differ from the teacher grading?
    a. Considering the actual grades awarded
    b. Considering the perceived value to the students of the comments made

Research Design

This study relates to a group case study assignment as part of a module on Supply Chain Operations. The module is a core component of a Master's Degree in Supply Chain Management. The feedback on the case study report consisted of written comments and an indication of the (provisional) grade achieved by the group. The rubric used, with typical comments, is shown in Appendix 1. The formative assessment on the report is important as part of a continuous learning loop where early feedback from this first report informs students' efforts on further continuous assessment exercises and the final examination.

The peer assessment exercise was designed based on the Reinholz (2015) assessment model with the following steps:

The task set was for each group to produce a report of their analysis of a case study on Supply Chain Operations. Each case was analysed by two or three groups of 5 students each, and a written report was submitted. (Task Engagement)

1. Each student was allocated a report to assess written by another group ('the target group') which had analysed the same case. Hence the students carrying out the assessment were familiar with the case and the analysis process. Each student assessed the case report individually using a preestablished summative grading rubric and provided written comments under each rubric element as well as a grade (See Appendix 1). (Peer Analysis)
2. Following the individual assessment, each group of students came together to compare their assessments of the target group report, and to generate a single integrated assessment again using the rubric. (Peer Analysis)
3. The synthesised rubric was then given to the target group and discussed with them. (Feedback Provision)
4. The group of students were then given feedback on their own report by another group. (Feedback Reception)
5. The reports were assessed by the lecturer using the same rubric. A (provisional) grade along with comments were given to each group on their report, only after the previous steps had been completed.

Students already had experience of peer review from other modules and hence were familiar with the process and had achieved a level of competence in the exercise. In assessing the work of a group rather than that of an individual, the emotional element of being seen to judge another student is minimised (Cho and MacArthur, 2010).

Data Gathering

In order to examine the effectiveness of the new peer assessment activity, data were gathered throughout and after the exercise

- The individual and group rubrics were analysed to see how individuals' grades are integrated into the group grade.
- The group rubric for each report was compared with the teacher rubric for the report, again considering the grades given.
- A survey was carried out to gather the views of the students on the effectiveness of the exercise, and the relative value of each stage in the peer assessment process.

The survey was administered following the university's Human Research Ethics Committee (HREC) guidelines. To comply with the guidelines, the survey was administered by a colleague unconnected to the module, and data access was not granted to the researcher until after grading was complete. The survey items were based on the existing literature, primarily on Reinholz (2010) for the impact of assessment on student learning, Cho and MacArthur (2010) for the comparison of lecturer and peer feedback, and Fraile et al. (2017) on the value of the rubric to students.

Results and Discussion

The survey examined how students perceived each step of the assessment process as contributing to the 4 learning outcomes.

Research Question 1

The general perception among students was that all 5 process steps contributed to all learning outcomes, with scores ranging from 4.33 to 5.00 on a 1-6 Likert Scale (Table 1 below). The learning from communication of feedback with the other group was rated lower than the receipt of feedback from the lecturer, while the learning resulting from actually reviewing the piece of work was rated highest. No learning outcome scored particularly highly compared to the others.

We then carried out an exploratory factor analysis of the learning items, to establish whether there were any patterns in the learning mechanisms at work. The results indicated that students did not strongly distinguish between the four learning outcomes – in general were students found an activity valuable, it was valuable across most if not all four learning outcome dimensions. The two strongest latent variables which emerged from this factor analysis are shown shaded in Table 1 below (orange and green), and they account for 58% of the variance in the results. What this suggests is that there are two underlying learning mechanisms at play – learning from receiving lecturer feedback and learning from giving and receiving peer feedback. The learning from carrying out the review of the work and from producing a combined group assessment was still strong, but did not evidence any particular pattern among the students. This suggests that individual students do learn from these first two steps, but in very different ways. In contrast, they tend to have similar learning patterns in steps 3/4 and step 5

Table 1 – Mean Learning Scores

Descriptive Statistics	it made me more aware of the assessment goals for this assignment	it made me more aware of how my group performed on this assignment	it helped me see how my group would be able to improve our performance on this assignment if given the opportunity	it helped me see how I could improve my performance on future assignments	Mean
1: When I carried out the review by myself of another group's report,	5.23	5.00	4.92	4.85	5.00
2: When my group had a discussion and produced a combined review of another group's report,	4.69	5.00	4.62	4.92	4.81
3: When my group delivered the feedback to another group on their report,	4.85	4.77	4.62	4.46	4.68
4: When my group received the feedback from another group on our report,	4.38	4.62	4.31	4.33	4.41
5: When I received the feedback from the lecturer on our report,	4.92	4.85	4.92	5.00	4.92
Mean	4.81	4.85	4.68	4.71	

Research Question 2a

We analysed the pattern of grades supplied in the assessment of each report, and compared the individual grades with the group grade (agreed by the group members in consultation) and with the lecturer grade. Even with the relatively small number of assessment groups, there was a clear pattern where the greatest consistency between individual, group and lecturer grades were for the middle-of-the road reports, which were typically a B+. For the reports where the lecturer grade was low (B-), the student grades tended to vary more within the group, and generally the students awarded a higher grade than the lecturer. On the other hand, where the lecturer grade was high (A-) there was again divergence in the student grades awarded, but this time students typically awarded grades lower than the lecturer's.

This pattern suggests that students have difficulty in assessing work that is either at the high end or the low end of the scale, and hence exhibit disagreement in the grades awarded. In addition, when their individual views are combined into a consensus grade, it tends to be closer to the class mean than the lecturer grade is. This suggests a reluctance among students to use the full range of grade classifications available. With work that is closer to the mean standard (assuming that the lecturer grade represents a valid indicator of the standard!) there was less disagreement among students as to the grade earned.

Research Question 2b

The comparison of students' perception of the value of the feedback comments shows that lecturer comments were more highly valued on all dimensions, as seen in Table 2 below. In most cases the differences between the value scores was statistically insignificant, except for the second question "The comments provided by *** on our report were well phrased and understandable". This finding runs counter to the view in the literature that peer comments tend to be more easily understood that lecturer comments due to the lack of academic distance involved (Cho and MacArthur, 2010). In this case, the lower rating across the board for peer feedback suggests that more work needs to be done in training the students how to assess and how to provide clear and relevant feedback comments.

Table 2 – Comment Quality Comparison

The comments provided by ****	Another Group	Lecturer
were specific and directly related to the assignment	4.15	4.92
were well phrased and understandable	3.85	5.08
will be of value in improving our performance in future assignments	4.38	5.00
were fair	4.38	4.92

Conclusions

The findings of this study suggest that all stages of the peer assessment process – task engagement, peer analysis, feedback provision and feedback receipt – contribute to student learning both in the short and long-term. Significant learning takes place in the task engagement and assessment discussion stages, but despite this the student perception is that the lecturer feedback stage is the most valuable form a learning perspective. This raises questions as to whether the lecturer involvement in the peer assessment process is providing added learning or is actually detracting from the peer-to-peer learning. The position of power that the lecturer holds as the arbiter of summative assessment may lead students to perceive lecturer feedback as more important than anything peers say.

The high level of variability in the peer grading indicates that further training of students is needed in the skills involved in assessing according to a rubric and effectively delivering and receiving feedback. It would be expected that this would reduce the variability in the peer grades and (hopefully) bring the peer grades more in line with the lecturer grades. This is particularly relevant for work at both ends of the quality spectrum.

Finally, it is interesting to see that students value the learning in terms of its impact on future assignments as as well as on the assignment under assessment. This supports the view that peer assessment builds evaluative skills which facilitate lifelong learning (Carless, 2007).

References

Ambrose, E. (2017) 'Getting Serious About Gaming', Proceedings of the 23rd International Annual EurOMA Conference, Trondheim, Norway.

Boud, D. and Falchikov, N. (2006) Aligning assessment with long-term learning. Assessment & Evaluation in Higher Education, 31 (4) pp. 399-413.

Boud D. and Falchikov, N. (2007) Developing Assessment for Informing Judgement. In D. Boud and N. Falchikov (Ed.), Rethinking Assessment in Higher Education (181-197) New York, Routledge.

Carless, D. (2007) Learning-oriented assessment: conceptual bases and practical implications. Innovations in Education and Teaching International, 44 (1) pp. 57-66.

Cho, K. and MacArthur, C. (2010) Student revision with peer and expert reviewing: Unravelling Peer Assessment. Learning and Instruction, 20 (4) pp.328-338.

Falchikov, N. (1995) Improving Feedback To and From Students. In P. Knight (Ed.), Assessment for Learning in Higher Education (157-166) Abingdon, RoutledgeFalmer.

Fraile, J., Panadero, E. and Pardo, R. (2017) Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. Studies in Educational Evaluation, 53 (Supplement C) pp.69-76.

Gibbs, G. and Simpson, C. (2004-5) Conditions Under Which Assessment Supports Students' Learning. Learning and Teaching in Higher Education, 1 pp. 3-31.

Grasha, A.F. (1994) A Matter of Style: The Teacher as Expert, Formal Authority, Personal Model, Facilitator, and Delegator. College Teaching, 42 pp. 149-142.

Hernandez, R. (2012) Does continuous assessment in higher education support student learning? Higher Education, 64 pp.502-489.

Jonsson, A. and Svingby, G. (2007) The use of scoring rubrics: Reliability, validity and educational consequences. Educational Research Review, 2 (2) pp. 130-144.

Liu, N.-F. and Carless, D. (2006) Peer feedback: the learning element of peer assessment. Teaching in Higher Education. 11 (3) pp.279-290.

National Forum (2016). Enhancement Theme Assessment OF, FOR and AS Learning: Continuing the Debate and Creating a Focus. Forum Insight. Dublin: National Forum. http://www.teachingandlearning.ie/wp-content/uploads/2016/08/Insight-Assessment-OFFOR-and-AS-Learning.pdf.

Nicol, D. J. & Macfarlane-Dick, D. (2006) Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Studies in Higher Education, 31 (2) pp. 199-218.

Reinholz, D. (2015) The assessment cycle: a model for learning through peer assessment. Assessment & Evaluation in Higher Education, 42(2) pp. 301-315

Tseng, S.-C. and Tsai, C.-C. (2010) Taiwan college students' self-efficacy and motivation of learning in online peer assessment environments.

Appendix 1: Case Study Report Assessment Rubric

Group __##__ Case: ____########____ Submission Date: _######____

Supply Chain Operations Small Case Report Grading Rubric					
			C- to C+	B- to B+	A- to A+
			Standard	Good	Outstanding
Explanation of the problem • The problem and its operational implications are clearly explained and contextualized • The discussion of the problem is linked to the material in the book and other assigned readings • An understanding of why the company is successful or not	15%	Good introduction and summary of key issues. Sets up the context well for the analysis to follow. Some good references used in addition to textbook material.	Shows some understanding of the issues and their operational consequences.	Shows adequate understanding of the issues and their operational consequences	Shows superior understanding of the issues and their operational consequences.
Sound arguments; The arguments are laid out in a clear and persuading fashion • Logically organized • Persuasive and compelling arguments • Creativity	15%	Yes, a well-presented argument for the ranking of objectives First and last were clear but would have been more persuasive if you had differentiated a bit more in the middle e.g. dependability is 'fundamental', speed and flexibility both 'essential'. Many creative elements e.g. outsource other projects and packaging	Arguments that supported the Case Analysis were adequately framed and borderline persuasive.	Arguments that supported the Case Analysis were well framed and persuasive.	Arguments that supported the Case Analysis were well framed and showed a deep understanding of the issues, resulting in an extremely persuasive write-up.
Appropriate analysis, given the information in the case • Use of financial and other data in the case to provide insightful analyses • Use of the case information to analyse "soft" aspects of the decision problem	25%	Soft data used very well, but financial payback data not used. Some repetition between Issues and Factors, but generally many very good points developed. Information security comes out as a major issue influencing the decision – should it have appeared in Objectives?	Adequate analysis that provides some insight on the decision	Well thought-out analysis that provides a solid basis for a decision	Insightful analysis that provides a deep and non-obvious outlook on the problem
Conclusions and recommendations are congruent with strategic analysis • Recommendations flow as a consequence of the analysis • Recommendations are actionable and concrete • Creativity of solution	25%	Yes, well developed and they tie in clearly with Issues/Factors. Strategy is concrete and feasible and presents Monica with a strong basis for persuading management.	Specific recommendations and/or plans of action provided are adequate. Specific aspects of the analyses are only loosely referred to in the conclusions.	Recommendations and/or plans of action are effective and well thought out. Recommendations build on the analysis and refer to specific results as a justification.	Recommendations and/or plans of action are creative and provide a great solution to the problem. Recommendations build on the analysis by creating a coherent and persuasive whole.
Proper organization, professional writing, and formatting • Proper formatting and structure • Grammar, spelling, and syntax	20%	Generally excellent. Word count well used.	Grammar, spelling, punctuation, professional writing, and syntax needs improvement	Adequate grammar, spelling, punctuation, professional writing, and syntax	Excellent grammar, spelling, punctuation, professional writing, and syntax