Student Engagement with Peer Assessment: a Review of Pedagogical Design and Technologies

Karen Stepanyan[1], Richard Mather[1], Hamilton Jones[1], and Carlo Lusuardi[1]

[1] School of Computer and Advanced Technologies, Faculty of Enterprise and Innovation, Buckinghamshire New University,
Queen Alexandra Road, High Wycombe, HP11 2JZ, United Kingdom
{Karen.Stepanyan, Richard.Mather, C.Hamilton-Jones, Carlo.Lusuardi}@bucks.ac.uk

Abstract. The paper summarizes participatory action research that explores student attitudes towards a peer assessment exercise and further reveals a distinctive pattern in student responses. A formative and reciprocal peer assessment exercise was studied to identify possible reasons for low levels of student participation. The target group included students in an undergraduate course in computing. A follow-up questionnaire, undertaken by 36 students, was analyzed and compared against assignment marks. Finally, the access statistics of the virtual learning environment (VLE) were examined. The major results indicate the following: [i] an expectation of more explanatory and supportive tutor intervention; [ii] a student preference towards anonymity; [iii] student interest in accessing peer work; and [iv] that the allocation of marks and in-class activities factors are important in encouraging student involvement.

Keywords: Peer Assessment, Peer Evaluation, Student Engagement, Formative Assessment, Educational Technologies

1 Introduction

Peer assessment is widely recognized to be the formative feedback and summative grading of individuals by peers of similar status [1]. Introduced as an innovative form of assessment, peer assessment aims at enhancing learning experience, assisting deep learning and enhancing the acquisition of critical thinking skills [2, 3]. Similarly to peer review process, used for evaluating work quality in professional occupations [4, 5], peer assessment encourages students to develop skills for the analysis and critical evaluation [6, 7].

This paper considers a case study built around a peer assessment exercise for assessing group work by undergraduate students. It aims to reveal potential vulnerabilities of methods and technical tools used for organizing and implementing web based peer assessment activities. Results from the study are interpreted to suggest measures for enhancing student engagement and the effectiveness of peer assessment.

2 Theoretical Background

A review of the literature indicates great variation in the models of peer assessment used in higher education. While originally used in writing courses [8], studies on peer assessment now span many subject areas.

Topping's review and typology of peer-assessment [9], records 17 models classified according to variations in characteristics such as outputs, privacy, official weight (summative contribution), participant ability. Topping also highlights the importance of considering anonymity when designing peer assessment exercises.

Earlier studies suggest that peer assessment may encourage students to engage in cognitively demanding activities. Examples include comparing, clarifying, contrasting, diagnosing, considering deviations and summarizing information. These activities are believed to reinforce knowledge and lead to better understanding and deeper learning [10]. Additionally, peer assessment supports development of teamwork and communication skills [11], and improves the understanding of institutional assessment processes [12].

3 Case Study

The study summarized in this paper is based on year-long modules concerning the development of database applications. The modules were offered to second year undergraduate B.Sc. and Fd.Sc. students in computing.

The peer assessment exercise was introduced as an optional part of a required and graded group assignment, for which 20% of the overall module marks were available. The assignment required students to work collaboratively in pairs, to design and develop a database. A written report was then submitted for marking. As part of the peer assessment exercise, students were asked to post specific sections of their reports for assessment by colleagues.

The typology of this exercise [13] could be expressed as being one of a formative, out of class, mutual, distance, not graded, voluntary, cross-ability and group peer assessment. The main incentive for student participation was an opportunity to improve work (consequently, grades) on the basis of suggestions made by their peers.

The peer assessment exercise was delivered in asynchronous mode using a discussion board on a Blackboard™ virtual learning environment. Each discussion board thread comprised the original report and the peer-reviews for each report.

The peer assessment exercise consisted of two tasks: [a] posting their group-work on the VLE; and [b] posting constructive feedback on the work of other groups. Students were allowed two weeks for completing the task. The exercise was thoroughly explained and tutor support was made available to the students throughout the peer assessment period. Blackboard announcements were posted and email reminders sent to all the participants to notify them of approaching deadlines.

Only four students in two groups completed the first peer-review task. Although posts were of high quality, the low level of participation was of some concern. Participant attitudes and behaviour were therefore investigated further with respect to: [a] critical reflection; [b] extent of passive (lurking) and active participation; and [c] by extending the study in an attempt to understand attitudes towards specific components of the peer assessment exercise.

4 Research Aims and Analysis

This investigation aims to understand attitude, behavioural, teaching and technical factors that may influence levels of student participation in peer-review exercises. Towards this end the afore-mentioned peer-assessment case study is assessed with reference to VLE log metrics and a 21-item questionnaire returned by 36 respondents.

The demographics of respondents are summarized in Table 1.

Table 1. Participant demographics

	Category	*Frequency*	*Percent*
Age	*19-20*	8	22.2%
	21-22	15	41.7%
	Over 23	12	33.3%
	Not Spec	1	2.8%
Sex	*Female*	8	22.2%
	Male	27	75.0%
	Not Spec	1	2.8%
Total	-	36	100%

4.1 Analysis of the Blackboard™ Access Records

In addition to active participation, log entries also contain records of passive presence (lurking) around the discussion, assessment and announcement areas established to support the peer-review process. Logs record 168 ‘views’ of posted materials by 18 students (50% of the cohort) accessing exhibited work and feedback. Log statistics therefore suggest: [a] a high level of interest amongst ‘passive’ participants in work submitted by colleagues; and [b] that passive “non-posting” involvement was much more widespread than active participation.

Despite seemingly low interest in initially posting contributions, the logs show that the peer assessment area continued to be accessed by students some time after the end of the exercise. Some 20% of all ‘hits’ recorded occurred up to two months following peer assessments with a further 6% logged during a 10 day period before students were to site a ‘time-constrained assessment’ (a formal test). Because this examination included problems similar to those given during the peer-assessment exercise, it seems very likely that students visited the peer assessment area for revision purposes.

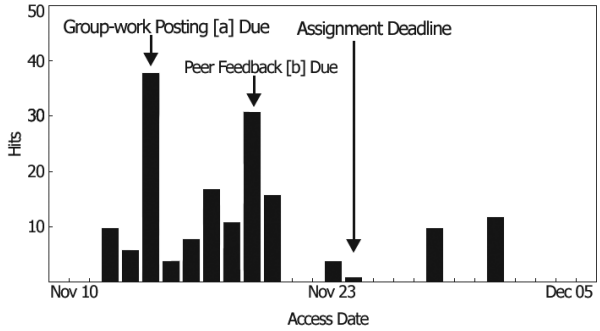

Fig. 1. Access to peer assessment area on Blackboard™ with deadlines indicated.

Results indicate that many students who did not actively participate did in fact 'passively' view content. A further cycle of study was undertaken to investigate possible reasons for low levels of active involvement, thereby suggest how greater participation in peer assessment work might be encouraged.

4.2 Student Feedback on Peer Assessment

A 21-item questionnaire was issued to determine student opinion concerning: [a] the rationale for peer assessment; [b] the design and delivery of the exercise; [c] levels of comfort/acceptance associated with elements of the peer assessment process; and [d] web technologies used for the exercise.

The great majority of students indicated that the exercise was fairly well explained and presented (86% or responses were recorded for categories of "satisfactory" and "very clear"). Additionally, results indicated that 67% of all respondents were interested in being able to view the work of their peers; this observation is also consistent with behaviour recorded in access logs. A greater proportion (78%) believed that the exercise could be beneficial. The proportion of those considering the exercise to be of no benefit (22%) was great enough to be of concern to the teaching team. While most students were interested in accessing their peer's submissions, only 50% were interested in providing feedback to their classmates.

The discovery that students were more inclined (78%) to engage if marks were awarded for participation is consistent with earlier studies [10]. A significant proportion (22%) suggested that one principal area of improvement would be to reward the quality/level of participation in peer assessments through summative grading.

The timing of the exercise was another factor shown to be important for increasing levels of participation. Two thirds of respondents indicated that timing would affect their level of engagement in peer assessments. Many preferred to conduct peer-reviews in-class rather than off-site and three students (8%) were particularly emphatic on this point (see Figure 2.).

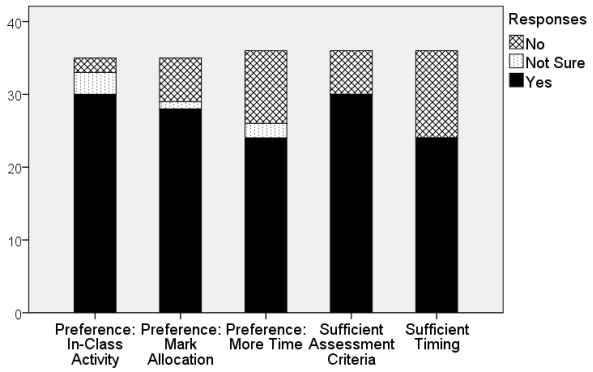

Fig. 2. Preferences in peer assessment exercise format.

Earlier research suggests that less desirable effects of peer-assessments may include increased participant workloads and anxiety levels [9]. The assessment of student 'comfort' level with regards to anonymity and workload revealed that 43% of participants were "very/uncomfortable" to post their work publicly. Students felt relatively more comfortable in terms of workload: 74.2% (mean=3.09, std. dev.= 0.951, n=35) of respondents indicated workload to be from "moderate" to "insignificant" (Figure 3). However, only 8.3% indicated insignificant increase in their workload, showing the demanding nature of the exercise.

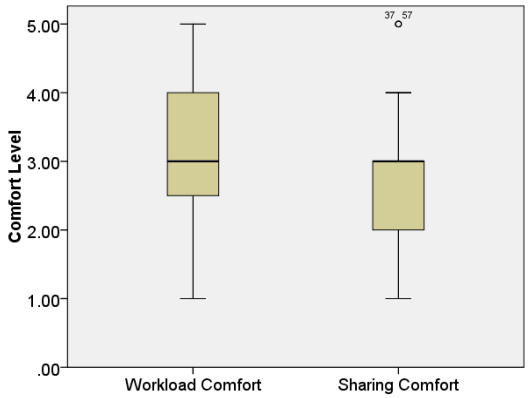

Fig. 3. Comfort level of students with [a] added workload and [b] required sharing of work (1: least; 5:most comfortable). Note: box plots are paired here for convenience of presentation only. Lickert scales for the two boxplot distributions should not be compared because they are nominal and represent different but (statistically) interdependent 'comfort' metrics.

For the purposes of triangulating/confirming the results reported above, students were also asked to express which improvement they believed might have the greatest

"participation encouraging" impact in peer assessment exercises. Results (see Table 1.) indicate that allocation of marks (22%) and clearer justification for exercises with provision of further support in completing reviews (together 20%) were the most important factors in influencing engagement with peer assessment. There was also notable support (11%) for delivering peer assessment exercises in-class.

Table 2. Student perceptions of the most important factors for encouraging participation in peer assessment.

Student Suggestions	*Freq.*	*%*
Clear explanation and justification of peer assessment	*2*	*5.6*
Clear explanation and more support	*5*	*13.9*
Active participation of others	*2*	*5.6*
Allocation of Marks	*8*	*22.2*
In-Class Activity	*4*	*11.0*
More Time Allocation	*2*	*5.6*
No Suggestions Made	*13*	*36.1*
Total	***36***	***100***

Comments on the Web tool. The VLE and discussion board in particular, were central in facilitating the peer assessment. A Blackboard™ environment that contained an announcement, resource and communication area, was made available for the purpose of the exercise. Discussion board threads were created for posted work and peer feedback.

VLE access records were notably consistent with questionnaire data; both sources indicated that 56% of all participants accessed the peer assessment area. Most of those who used the system (91%) rated access as being 'very' to 'moderately' easy. However, because 44% of participants did not visit the VLE, this observation cannot be extrapolated to the remainder of the cohort.

4.2 Correlation Analysis of Survey Data against Student Marks

Student marks for the first two module assignments were averaged and were then analyzed alongside questionnaire data. The intention of analysis was to reveal patterns between responses and mark performance. For example it was of interest to know if higher achievers were more likely to appreciate the benefits of peer assessment or more willing to provide feedback. Due to the non-parametric nature of the data Kendall's tau statistic was used for calculating correlation coefficients (Table 3).

The results indicate that average marks (AM) and anticipated level of additional workload (ALAW) associated with peer assessment are significantly correlated (r=-.335). The direction of this relationship suggests that higher achieving students are more likely to report lower level of expected extra workload as a result of undertaking peer assessment. The statistically significant correlation (r=.334) between student average marks (AM) and willingness to access peers work (WAPW) suggests that higher achieving students are also more likely to be interested to see the work of their

peers. While some practitioners may expect that lower achievers would be most interested in work of others, the results suggest the opposite.

Table 3: Bivariate Non-parametric Correlation (Kendall's tau) of Collected Questionnaire Data and Student Marks.

Kendall's tau_b		*CI*	*AU*	*ALAW*	*WTF*	*WAPW*	*AM*
CI (Clarity of Introduction)	Cor. Coeff.	1					
	Sig. (2-tailed)	.					
	N	36					
AU (Anticipated Usefulness)	Cor. Coeff.	0.205	1				
	Sig. (2-tailed)	0.154	.				
	N	36	36				
ALAW (Anticipated Level of Additional Workload)	Cor. Coeff.	-.431**	0.038	1			
	Sig. (2-tailed)	0.003	0.795	.			
	N	35	35	35			
WTF (Willingness to Receive Feedback)	Cor. Coeff.	-0.025	0.13	0.187	1		
	Sig. (2-tailed)	0.876	0.408	0.246	.		
	N	35	35	34	35		
WAPW (Willingness to Access Peers Work	Cor. Coeff.	0.252	0.071	0.18	.356*	1	
	Sig. (2-tailed)	0.112	0.649	0.257	0.041	.	
	N	35	35	35	34	35	
AM (Average Mark TCA&CW1)	Cor. Coeff.	0.076	-0.002	-.335*	0.181	.334*	1
	Sig. (2-tailed)	0.572	0.987	0.014	0.219	0.024	.
	N	34	34	33	33	33	34

Notes: [1] ** = Correlation is significant at the 0.01 level (2-tailed).
[2] * = Correlation is significant at the 0.05 level (2-tailed).
[3] CI, AU and ALAW are likert scale, and WTF and WAPW are dichotomous (y/n) data

Another, less surprising result, indicates a highly significant correlation (r=-.431) between anticipated level of additional workload (ALAW) and the clarity of introducing (CI) the peer assessment exercise. The results suggest that students are more likely to record lower level of anticipated workload when clarity of introduction is reported to be greater. No other statistically significant correlation was discovered. As in any correlation analysis these findings do not imply causality. Yet, these results may guide practitioners when facilitating such exercises.

5 Use of VLE for Peer Assessment

The VLE used in this study had no dedicated facility for peer assessment. Group management and discussion boards can be used as an alternative, but they impose certain limitations. A group management facility provides closed communication tools, such as chat, forum and group emailing. These tools can be used for collaboratively completing tasks before submitting them for peer assessment. However, the issues identified in the questionnaire show that alternative technologies, for example discussion boards, may lack certain desirable features. Such features

include those that ensure confidentiality but allow facilitators to identify individuals participating in assessments. However, the use of such additional tools may discourage participation since these will involve familiarization and increased exercise workloads. Thus, any perceived benefits should be carefully balanced against the additional learning burden of adopting new technologies.

6 Conclusion

The results suggest that students, in one way or another, were largely positive about the peer assessment process. However, some critical issues that deserve the attention of practitioners and researchers were identified. These included that: [i] students may not initially perceive the rationale for and potential benefits of peer assessment. Nevertheless, average marks were not found to be correlated with perceptions of 'anticipated usefulness' of peer assessment; [ii] student marks were significantly correlated (r=-.335) with anticipated levels of increased workload, suggesting that peer assessment activities must be designed to meet the needs of all students; [iii] from the questionnaire study it was apparent that participants were very interested in the solutions submitted by their colleagues. Correlation analysis suggested that those achieving highest marks were most interested in studying the solutions of their colleagues (r=.334). The results of the VLE access statistics are consistent with this finding and indicated that solutions posted were viewed long after the peer assessment exercise had been completed, possibly for purposes of revision; [iv] many students are not completely comfortable with posting work publicly, often preferring to remain anonymous when making their own submissions and when assessing peers; and [v] questionnaire results suggest that grading and in-class work are the very important factors for encouraging participation.

Overall, findings strongly suggest that design, delivery techniques, facilitation methods and specific features of technology (for example, those that allow a degree of anonymity) are important factors for creating 'safe' learning areas that encourage greater participation in peer-assessment exercises.

6 References

1. S. Bostock, "Student Peer Assessment," *Keele University, Learning Technology, Documents, December,* 2000.
2. G. Mowl, "Innovative Student Assessment," in *Deliberations*, 1996.
3. L. McDowell, "The Impact of Innovative Assessment on Student Learning," *Innovations in Education and Teaching International,* vol. 32, pp. 302-313, 1995.
4. S. Cole and G. A. Simon, "Chance and consensus in peer review," *Science,* vol. 214, pp. 881-886, 1981.
5. F. E. Bloom, "The Importance of Reviewers," *Science,* vol. 283, p. 789, 1999.
6. J. Liu, D. T. Pysarchik, and W. W. Taylor, "Peer Review in the Classroom," *BioScience,* vol. 52, pp. 824-829, 2002.

7. D. W. Johnson, R. T. Johnson, and K. A. Smith, "Active Learning: Cooperation in the College Classroom," 1998.
8. G. Jacobs, "First Experiences with Peer Feedback on Compositions: Student and Teacher Reaction," *System,* vol. 15, pp. 325-33, 1987.
9. K. Topping, "Peer Assessment Between Students in Colleges and Universities," *Review of Educational Research,* vol. 68, p. 249, 1998.
10. K. A. Van Lehn, M. T. H. Chi, W. Baggett, and R. C. Murray, "Progress report: Towards a theory of learning during tutoring," *Pittsburgh, PA: Learning Reasearch and Development Center, University of Pittsburgh,* 1995.
11. S. M. Riley, "Peer Response in an ESL Writing Class: Student Interaction and Subsequent Draft Revision," Florida State University, 1995.
12. S. A. Fry, "Implementation and Evaluation of Peer Marking in Higher Education," *Assessment & Evaluation in Higher Education,* vol. 15, pp. 177-189, 1990.
13. K. J. Topping, E. F. Smith, I. Swanson, and A. Elliot, "Formative Peer Assessment of Academic Writing Between Postgraduate Students," *Assessment & Evaluation in Higher Education,* vol. 25, pp. 149-169, 2000.
14. R. Ballantyne, K. Hughes, and A. Mylonas, "Developing Procedures for Implementing Peer Assessment in Large Classes Using an Action Research Process," *Assessment & Evaluation in Higher Education,* vol. 27, pp. 427-441, 2002.