

Student-generated content: using PeerWise to enhance engagement and outcomes in introductory physics courses

Simon P. Bates, Ross K. Galloway and Karon L. McBride

*Physics Education Research Group, School of Physics and Astronomy,
University of Edinburgh, Edinburgh EH9 3JZ, UK*

Abstract. We describe the implementation and evaluation of an online tool to support student generation of multiple choice assessment questions within two consecutive semesters of introductory physics at the University of Edinburgh. We substituted a weekly homework for an assessment activity in which each student was required to participate in using the system. Engagement with the system was high, with contributions generally going beyond the minimum requirements. The quality of submissions was on average high, with the very best questions being remarkably detailed problems rather than exercises. We explore links between use of the online system and end of course examination score. We find that students with higher levels of activity in the system scored significantly higher marks on the exam; this effect was seen for students of lower ability as well as for the highest performing students.

Keywords: assessment, student-generated content, e-learning

PACS: 01.40.Di, 01.40.Fk, 01.40.gb

INTRODUCTION

Advanced and refined problem solving skills are rightly prized as one of the key attributes acquired during a physics degree. To this end, by the time our students graduate, they have probably answered many hundreds of physics problems set for them by instructors. They are unlikely to have ever authored any themselves. This paper investigates using student-generated assessment content (SGC) in a pair of consecutive introductory physics courses. The central idea is to place the students in the role of active creators of content in their courses, rather than merely passive consumers of static content prepared by instructors. SGC has been advocated [1] as a means to foster deep learning and high levels of student engagement, leading to enhanced conceptual understanding. In a sense, it is a representation of the adage that many of us will be familiar with, that you “don’t really understand something until you have to teach it”.

We have employed the PeerWise online system [2, 3] that allows students to author their own multiple choice assessment questions (and associated explanations), as well as answer those created by their peers, rate questions, provide comments on questions or answers, and seek help from question authors. It incorporates social functionality that is common amongst Web2.0 applications (such as the ability to ‘follow’ favourite question authors) and provides significant opportunities for self- and peer-assessment and timely formative feedback [4, 5].

PeerWise has been deployed in a wide range of undergraduate courses around the world, with detailed analy-

ses of student engagement and performance in Computer Science and Engineering courses having been reported [6, 7]. Here we follow the presentation and analysis reported in Denny *et al.* [8] to investigate levels of engagement with the system and to probe correlations between use and end of course outcome on the final exam. This is also one of the first studies to examine use of PeerWise in a heavily mathematical subject.

IMPLEMENTATION

The first year physics courses at the University of Edinburgh (Physics 1A and 1B) are taken by 200-300 students per year. The courses run in consecutive semesters with almost all Physics 1B students having taken 1A in the previous semester. In both courses, approximately half the cohort are studying towards a physics degree, whilst the other half are reading other subjects (usually science-based, but not exclusively). All students taking the course are required to have studied physics to the final year of high school. The courses have developed over several years to incorporate significant elements to foster interactive engagement: use of clickers in lectures, supplemented by workshop (studio) tutorial classes.

PeerWise was incorporated into the assessment structure for the two courses in place of a standard weekly homework exercise. Students were required to contribute a minimum of the following: one question authored, five questions answered, three questions commented on and rated for quality and difficulty. In Physics 1A, the assessment deadline was placed one week after the task was

set; in Physics 1B it was at the end of the course, some 7 weeks later. The assessment counted for approximately 3% of total course credit in Physics 1A and 1% in Physics 1B. Assessment marks were derived from the scoring system in PeerWise, which allocates points to students based on the number and quality of questions authored, answered, commented on, etc.

Prior to the first exercise in Physics 1A, we devoted approximately 1 hour of the 3-hour weekly workshop to providing support and scaffolding to students in terms of what makes a good MCQ, and how to go about devising questions that were just beyond their current understanding of the material. This scaffolding took the form of a structured group work activity with a worked example question prepared using a template to guide student thinking, followed by an exercise in which groups (5-6 students) devised a question using the same blank template. We did not spend time showing students how to use the system and instead provided short screencasts of common functionality queries (such as how to author questions, how to find questions on a certain topic, etc.). These screencasts have been made available on the PeerWise web site [2]. During the period in which the PeerWise assessment tasks were live, we did not intervene in any way, but merely observed the growth of the question repository and discussions that took place within it. On more than one occasion we observed an incorrect explanation and / or solution to a problem rapidly corrected by a peer.

EVALUATION

During the execution of the two assessments using PeerWise, students contributed nearly 700 original questions. A full and detailed analysis of the types of question, comments and patterns of student use is on-going. In the analysis presented below, we focus on overview measures of student engagement followed by an investigation into correlation of student use of PeerWise with end of course examination mark. We follow the methodology reported by Denny *et al.* [8], in which the cohort is split according to their level of activity in PeerWise into those higher / lower than the median activity for the assessment. We further sub-divide the cohort into quartiles of initial ability (based on performance on the FCI prior to the start of the course) and re-perform the same median split analysis.

Engagement

The initial scaffolding activity presented to students ahead of the first PeerWise assessment in Physics 1A

TABLE 1. Summary data for student contributed questions, answers and comments by course.

	Questions	Answers	Comments
1A (N=197)			
Total	347	3258	1903
Maximum	17	145	125
Mean*	1.8 ± 0.1	16.5 ± 1.4	9.7 ± 1.0
Median	1	11	6
1B (N=188)			
Total	346	4065	2314
Maximum	49	338	303
Mean	1.8 ± 0.3	21.6 ± 3.3	12.3 ± 2.3
Median	1	7	4

* with standard error

was very well received by students. There was a notable ‘buzz’ in the workshops when students were engaged in this activity and it was difficult to get them to move on to the next part of the programme of activities at the end of the allocated time for the PeerWise introduction. This initial enthusiasm carried forward into use of the system during the assessment period (which for Physics 1A lasted 1 week). Table 1 shows quantitative summary data for contributions by students in each of the two Physics courses. In Physics 1A, 95% of the total cohort engaged with the Peerwise activity, and this dropped to 63% in Physics 1B. (This may be a reflection of that fact that we vigorously and repeatedly encouraged the Physics 1A students to participate, but were more passive for Physics 1B.) Although there was a large spread in the number of questions, answers and comments contributed, all mean values exceeded the minimum stipulations for the assessment (which were one question written, five questions answered and three comments and ratings made, as described above). The figures for the second semester Physics 1B course show approximately the same number of student contributed questions, but a higher number of both answers and comments. A smaller fraction of the class participated in the second semester activity (though it was undoubtedly the more engaged fraction of the whole class who did!) but the assessment was live for several weeks, as opposed to the initial one week deployment in Physics 1A. Although a fraction of the comments were empty of any physics discussion ("Great questions, thanks"), many comment threads showed evidence of real discussions about difficulties with problem solving and, more often than not, them being resolved.

Figure 1 shows a time-resolved picture of one of the values from Table 1: the number of contributed answers as a function of time for the Physics 1B course. The assessment deadline was 25th March 2011 and, unsurprisingly, student usage peaked around this time (though not as sharply as has been observed in some other studies,

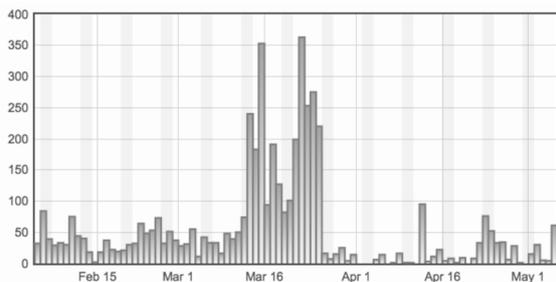


FIGURE 1. Number of contributed answers as a function of time for second semester (Physics 1B) course.

with an order of magnitude increase in activity immediately prior to a deadline!) Usage at a lower level continued during the revision period and up to the date of the final examination (6th May 2011) even though there are no MCQ questions on the final examination paper.

The quality of student contributions was very high, and in some cases, truly exceptional. A full analysis is not practicable here, but we note in passing that the overwhelming majority of questions contributed during the first assessment focussed on quantitative problem solving (usually with a context, often humorous). In the second semester, there was a greater proportion of more conceptual questions. This may be a consequence of the content of the respective courses: Physics 1A is a mechanics course, and the students seemed able to invent richly contextualised questions based on real-world situations; in contrast, Physics 1B focusses more on modern physics, a context in which the students seemed less inclined to set extended, quantitative problems. In both assignments, students made extensive use of mathematical notation, usually (but not always) through the built-in equation editor interface. Students also made extensive use of images and diagrams in their questions (and answers and comments). These were produced using a variety of technical solutions, from what looked like sketching programs, through standard figure-drawing packages, to digital photographs of hand-drawn sketches on paper.

The students were asked to provide their views of PeerWise in the end-of-course evaluation questionnaire. Whilst we should observe some caution in interpreting student opinion before the effects (or otherwise) of the system have had time to be fully absorbed by them, nevertheless their views were broadly positive. The students were in general agreement that writing questions improved their understanding of the course topics; opinions were less strong on the benefits of answering other people's questions, but were generally positive. Written comments supported our subjective impression of PeerWise having increased levels of engagement and enthusiasm in the course.

Attainment

Using the methodology of Denny *et al.* [8], we have investigated the way in which use and engagement with PeerWise may be correlated with end of course performance. Here, we are careful to separate out correlation and causation: we are only able to seek evidence for the former and cannot isolate the latter. We performed a median split of the class for the PeerWise assessments on the basis of activity on the system, separating students into categories of High PeerWise Activity (HPA) and Low PeerWise Activity (LPA), for values that were above / below the median value for that assignment. We then looked to see if there were any differences in mean mark on the end of course examination for these cohorts. The results for both courses are shown in Table 2.

Table 2 illustrates that there is a statistically significant difference in the mean exam mark of students in the HPA sub-cohort, compared with those in the LPA group, for the courses in both semesters.

It may be the case that this is simply evidence of better students performing more highly, whichever type of assessment we happen to set for them. To investigate this, we have pre-divided the cohort into ability quartiles, as measured by performance on a conceptual test of mechanics understanding (FCI) taken by students at the start of the course. We obtain an extra quartile, containing both those students who are not sufficiently organised to take the test before the deadline and also students who join the course late. We then perform an equivalent median split for students within each of these quartiles, dividing members of each quartile into HPA and LPA groups, and compare with the mean exam mark.

Results for the Physics 1A course are shown in Figure 2. We find that the mean exam score of the HPA group is significantly higher than that for the LPA group within the same quartile for Q1, Q4, and those students without a pre-test score. Whilst the largest effect in Q1 may be related to this highest quartile containing a disproportionately large fraction of above average question authoring activity, it is clear from Figure 2 that even the students initially in the lowest quartile of ability (Q4) at the start of the course who engage with PeerWise effectively achieve better outcomes on the end of course examination. A similar analysis for the Physics 1B class differs in detail but reproduces the main finding: improved exam performance is not restricted to only the higher performing students but extends to the lowest quartile.

DISCUSSION AND CONCLUSIONS

We decided to pilot the introduction of PeerWise in these courses last year, substituting a single assessment in each course for a PeerWise activity. It seems apparent that in-

TABLE 2. Comparison of end of course exam mark with PeerWise activity level.

Cohort	Number of students	Mean exam score*	Standard error	<i>p</i> value	Effect size [†]
1A (<i>N</i> =193)					
HPA**	104	63.2	1.6	<0.001	0.29
LPA	89	53.6	1.6		
1B (<i>N</i> =182)					
HPA	94	61.9	1.8	<0.001	0.36
LPA	88	46.8	2.4		

* all scores expressed as percentages

[†] Pearson's *r*

** HPA / LPA denote higher / lower than median PeerWise activity

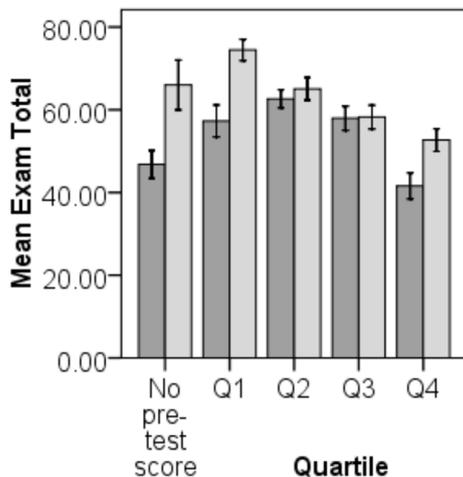


FIGURE 2. Physics 1A end of course examination mark for lower / higher than median PeerWise activity sub-cohorts (LPA, dark bars and HPA, light bars, respectively), split by pre-course FCI performance. Error bars represent \pm the standard error on the mean.

corporation into the (summative) assessment strategy of a course is essential to ensure engagement of a representative fraction of the class, not simply those who will eagerly tackle anything and everything that we ask of them. The quality of questions that were authored, and the discussions and interactions that followed from them, was influenced by the initial scaffolding activity that we undertook in class time ahead of the first PeerWise activity. We believe this had the effect of setting the bar at a high level, illustrating to students the desired type of questions that were sought, resulting in very small numbers of trivial or unrelated questions. A more detailed evaluation of the nature and type of questions contributed by authors, together with types and patterns of comments that students provide, is underway. Even on the basis of the analysis completed thus far, the pilot provides evidence of a rich vein of student creativity that has thus far

gone largely untapped within these (and other) courses. Additionally, it has provided valuable input to our re-evaluation of the assessment strategy in these courses for the coming academic session.

ACKNOWLEDGMENTS

We acknowledge the support and assistance of Paul Denny, author of the PeerWise system, and the Higher Education Academy UK Physical Sciences Centre for project funding.

REFERENCES

1. S. Draper, *British Journal of Educational Technology* **285**, 285–293 (2009), URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8535.2008.00920.x/pdf>.
2. (Accessed 4th July 2011), URL <http://peerwise.cs.auckland.ac.nz/>.
3. P. Denny, A. Luxton-Reilly, and J. Hamer, *Proceedings of the tenth conference on Australasian computing education* (2008), URL <http://portal.acm.org/citation.cfm?id=1379255>.
4. D. Macfarlane-Dick, *Studies in Higher Education* **31**, 199–218 (2006), URL <http://www.informaworld.com/index/743789423.pdf>.
5. D. Nicol, *Journal of Further and Higher Education* **31**, 53–64 (2007), URL <http://www.informaworld.com/index/772826527.pdf>.
6. P. Denny, J. Hamer, and A. Luxton-Reilly, *20th Australasian Association for Engineering Education Conference* (2009), URL <http://aaee.com.au/conferences/AAEE2009/PDF/AUTHOR/AE090080.PDF>.
7. P. Denny, B. Hanks, and B. Simon, *Proceedings of the 41st ACM technical symposium on Computer science education* (2010), URL <http://portal.acm.org/citation.cfm?id=1734407>.
8. P. Denny, J. Hamer, A. Luxton-Reilly, and H. Purchase, *Proceedings of the fourth international workshop on Computing education research* pp. 51–58 (2008).