Writing multiple-choice questions to access the higher levels of Bloom's taxonomy

Christopher J. Adams

School of Chemistry, University of Bristol, BRISTOL BS8 1TS, UK

Corresponding Author: chris.adams@bristol.ac.uk

Abstract

I report the implementation of an activity in which students are asked to write multiple-choice questions (MCQs) on the subject of 'orbitals' in order to consolidate their learning on the subject. This was facilitated using the online system PeerWise, which allows students to upload MCQs that they have written and to then answer those authored by their peers. The process of writing questions accesses the upper levels of Bloom's taxonomy, and the discussions incorporated within the activity allow for socially constructed learning as part of the pedagogy of constructive evaluation.

Introduction

Many experienced educators will probably agree that it is often much harder to write a good multiple-choice question than it is to answer one, and that a more thorough understanding of the subject matter is needed. This can be understood by reference to Bloom's taxonomy of learning: multiple choice questions can often be answered by remembering information, or applying a procedure, and remember and apply are verbs from the lower half of the revised version of the taxonomy. Conversely, in order to create a question the highest level of cognition is required (Krathwohl, 2002).

In recent years, educators have begun to exploit this in order to encourage deep learning in students, by creating assessment tasks that not only require students to answer questions, but also to write them. An early example was the work of Fellenz, who constructed an assessment regime that he called Multiple Choice Item Development Assignment (MCIDA) for business students (Fellenz, 2004). This required students to construct three sets of three multiple choice questions (MCQs), each consisting of a stem with one correct answer (the key) and three incorrect answers (distractors). The students were also required to identify the correct answer, providing an explanation as to why it was correct, and to explain why the distractors had been chosen and why they were not correct. Additionally, students were asked to identify which level of Bloom's taxonomy would be required to correctly answer the question. Between submission of each set of questions the students would receive instructor feedback on the quality of their questions, thus allowing them to improve subsequent attempts. The best questions were used as part of the summative end of module MCQ assessment, and the MCIDA contributed 20% of the course marks.

Fellenz reported a number of benefits from MCIDA, not the least of which was that it aided learning of the course content: in having to construct explanations and justifications of the key and the distractors, students are forced to spend time on task and to engage deeply with the subject matter. He also found that it improved student awareness of cognition and

how it can be improved, and of strategies for answering MCQs. The activity demonstrates how a well-constructed MCQ task can align with the seven principles of good feedback practice articulated by Nicol and Macfarlane-Dick (2007).

More recently, the use of MCQ writing assignments has been greatly facilitated by the PeerWise website (*PeerWise*, no date, b). Based at the University of Auckland, this is a free online tool to facilitate student construction, peer review, and answering of MCQs. Students author and upload questions and answers with explanations; other members of their class are then able to attempt these questions and leave ratings of difficulty and quality, and comments, and can flag questions they think may be wrong and praise questions they think are good. The system tracks participation and awards points and badges for taking part, introducing an element of reward and gamification (Ryan, 2013). Data generated by the site allows instructors to monitor student engagement.

PeerWise has been widely used across a number of disciplines, (PeerWise, no date, a) with a few studies reporting its use in chemistry. At Dublin Institute of Technology, PeerWise was used as the independent learning component in a first-year organic chemistry class, replacing recommended reading lists (Ryan, 2013). The unit assessment was also changed to be through MCQs, thus ensuring that the activity and the assessment were aligned. Increased student engagement and self-regulation was reported, with the social and competitive aspects of the system reported to be a powerful motivational factor.

At the University of Hertfordshire, Fergus has used it as a formative part of a foundation chemistry module for a cohort of first-year undergraduate pharmacy and pharmaceutical science students. The activity was introduced with a workshop that explored the structure of a good-quality MCQ using non-chemistry examples, and which then tasked small groups of students with creating their own MCQ, which was then evaluated by the group. Following the workshop, students were then required to individually construct two questions on subjects specified by the instructor, and then

add them to the PeerWise website. They then had to answer five of their peers' questions, and leave feedback on three. Evaluation of the activity revealed that 95% of the class successfully engaged with it, and that many more questions were answered and comments left than would be created by the bare minimum of participation. Students enjoyed the online community that the activity created, and reported that both writing and answering the questions improved their understanding of the subject matter (Fergus, 2019).

Similar encouraging conclusions were reported in studies from the Universities of Nottingham, Edinburgh and Glasgow (Casey *et al.*, 2014). Again, the activity was scaffolded by discussion of good and bad example questions, especially the need for good explanations and plausible distractors, and discussion of the rationale and pedagogy behind the exercise. Students were specifically asked to try and deepen their understanding by writing questions that targeted material in their 'zone of proximal development'. In these implementations a small amount of course credit (between 2% and 5%) was attached to the activity, but again student participation far exceeded the minimum required to gain that credit. A statistically significant correlation was found between academic performance and PeerWise engagement, with evidence that lower performing students benefitted the most (Hardy *et al.*, 2014) and it appears that time-on-task is the most useful indicator of benefit: students spending a lot of time on PeerWise do better in their exams than students who spend less time there (Galloway & Burns, 2015; Kay, Hardy & Galloway, 2020).

The particular pedagogy of learning through constructing, answering and evaluating assessment items has been called 'constructive evaluation' (Luxton-Reilly & Denny, 2010). As already discussed, authoring MCQs can lead to improvements in understanding; then, in submitting the answers and explanations to the questions they have written to the system, students are exposing their perspectives and thought processes to others and revealing how the factual content of the question fits into their knowledge structures. Dialogue with question solvers may reaffirm this, solidifying the author's confidence, or it may challenge it and force the

author to reconsider their beliefs. The same process operates for the solvers: the explanation received after answering questions might support the rationale that they had already constructed, but it may provide an alternative viewpoint for them to consider and assimilate. PeerWise thus provides an environment for students to socially construct and explain and defend their ideas, and through this to improve their conceptual understanding (King, 1990).

Implementation in Bristol
The current work reports the implementation of PeerWise at the University of Bristol in the 2020-21 academic year. It was used to support a short first-year module on the subject of orbitals. The cohort of 215 students had been introduced to atomic orbitals previously, and this course then introduces molecular orbitals and their role in simple organic reactions. It constitutes about one-eighth of the complete first-year introductory chemistry unit.

Because of the COVID-19 pandemic, the course was taught entirely online in a blended fashion. New content was introduced using short videos and screencasts of no more than 15 minutes duration, which the students could access at their convenience (asynchronously). These were supported using timetabled synchronous session held online using the Zoom videoconferencing software, which included interactive activities using Mentimeter and Padlet, and two problem classes (known to us as workshops) where the cohort was split into groups of 20-30 and worked through short-answer questions. There were also online computational activities to help visualisation, and a workbook that the students could complete.

We feel that the topic of orbitals is something of a threshold concept for students, and that it is critical that they fully grasp it in order to support subsequent learning (Talanquer, 2015). Therefore, the PeerWise activity was carried out at the end of the module, and was designed as much to encourage student learning of the material as to assess it. It was nonetheless given a weighting of 10% towards the final unit mark, as it is generally found that doing so encourages students to participate (Luxton-Reilly & Denny, 2010; Ryan, 2013;

Casey *et al.*, 2014; Hardy *et al.*, 2014; Mac Raighne *et al.*, 2015).

Previous reports of PeerWise almost unanimously stress the importance of properly scaffolding the activity (Luxton-Reilly & Denny, 2010; Devon *et al.*, 2011; Bates *et al.*, 2014). Fergus said 'This isn't a case of setting up PeerWise and expecting it to just work for your students on its own' (Fergus, 2019) – and therefore great care was taken over this. It was introduced using a short asynchronous video that explained the rationale for the activity and situated it within Bloom's taxonomy, and which discussed examples of questions that accessed the lower and the higher levels of that taxonomy. It also introduced the terminology associated with MCQs (such as key, stem, distractor) and highlighted some common pitfalls to avoid when writing MCQs, such as implausible distractors and confusing syntax. Finally, it introduced our definition of a 'good' question, which we define as one that requires more than simple factual recall to answer.

After viewing the video, students were then asked to complete two short MCQ tests. The first was designed to highlight more things to avoid when writing MCQs, and used the questions from the well-known content-free 'grunge-prowkers' test (Race, no date). The second contained examples of 'good' questions on the subject of orbitals, to illustrate the kind of question that we wanted students to produce.

The students were then given their first task: they were asked to construct a single 'good' MCQ, on the subject of orbitals. The question should have one correct answer, with an explanation, and three distractors, with explanations of how they are wrong and also why they were chosen. Students were also asked to indicate which level of Bloom's taxonomy they thought would be reached in answering the question. They were encouraged to try and write questions that lay within their zone of proximal development, so the activity would explore content and concepts that they found challenging.

Week 1	Week 2	Week 3	Week 4	Week 5
Asynchronous work: orbitals	Asynchronous work: orbitals	Asynchronous work: orbitals	Asynchronous work: MCQs	Asynchronous work: Peerwise
Synchronous sessions	Synchronous sessions	Synchronous sessions	Synchronous MCQ tutorial	
Workshop		Workshop		

Table 1 Outline of the structure of the orbitals course.

The questions were required for the fourth week of the course – that is, the week after lectures had finished – in which every student had a timetabled tutorial (Table 1). These were hour-long synchronous sessions in which small groups (typically 4-5 students) and an academic staff member met over Zoom. This tutorial was dedicated to the MCQ activity. Students were asked to bring along a question they had written, and they were subject to review during the session. The students attempted each other's questions, and discussed them, pointing out any ways in which they might be improved and whether or not they agreed with the author's estimation of the cognitive level required to solve it. This was all facilitated by the staff member present. A student preparing a question and participating in the tutorial was awarded half of the available credit.

There were two main reasons for introducing a tutorial into the activity before PeerWise was used. The first was to create an extra opportunity for constructive evaluation, by allowing the students to attempt the questions as a group of peers. It was hoped that this would surface any misconceptions, which could then be addressed either by peers or by tutors. In discussing whereabouts in Blooms taxonomy a question lay, students might have to consider how to answer the question, thus exposing them to alternative strategies and ways of thinking.

The second reason was to create confidence in the questions. A common student reservation about PeerWise is that the supplied answers may be incorrect (Mac Raighne *et al.*, 2015), and having them all vetted by tutors before being uploaded should prevent this. It is worth noting at this point that previous studies

into this have found that the incidence of wrong answers and explanations is low, and that these instances are normally picked up and flagged by participants (Luxton-Reilly & Denny, 2010; Hardy *et al.*, 2014) though a "robust review process during the development stage" of student-generated instructional material has been recommended (Coppola, 2015). Additionally, the tutorial should give students confidence in the quality of the questions (i.e. that they are 'good' questions) and should also weed out any obvious instances of plagiarism, another common student concern (Mac Raighne *et al.*, 2015).

Following the tutorial, the students were required to upload their question to PeerWise. They were then tasked with achieving a PeerWise reputation score of 1000 points, which can be achieved by authoring more questions, answering other students' questions, and commenting upon questions (Denny, no date). They were given a further week to do this (until the end of week 5), and successfully doing so would gain the remaining half of the available credit.

As this is the first time we have used this activity, we wished to evaluate its success (or otherwise). In particular, we wished to try and assess:

- Is a PeerWise reputation score of 1000 a suitable target, how many students reached it, and what strategies did they adopt to do so?

- What evidence is there for the activity improving the students' understanding of orbitals?

Question number	1	2	3	4	5	6	7	8	9
Strongly agree	23	21	7	24	14	42	16	1	8
Agree	33	34	14	23	32	12	19	8	17
Neither agree nor disagree	1	2	24	5	7	3	16	12	18
Disagree	0	1	10	2	2	0	6	32	13
Strongly disagree	0	0	2	3	2	1	0	3	2
Question number	10								
Quite a lot	1								
A bit	29								
Not at all	25								
I didn't take a question to my tutorial	3								

Table 2 Responses to the post-activity questions

Methods

The study reported herein uses three main sources of evidence. The first is that data held within the PeerWise system itself, that is, the questions, answers, explanations and comments left by the students as part of the exercise. As well as this, there is also the metadata associated with PeerWise, such as information about when and how many contributions were made.

The second sources of evidence are recordings of the tutorials that were held in week 4. The cohort was split into 46 smaller groups for this activity, and a random selection of four of the tutorials was recorded using the inbuilt functionality of Zoom, which was also used to produce a transcript of the tutorial. This was manually corrected by comparison with the recording, and then anonymised using letters of the alphabet for the students. All the recordings were made with the express permission of all the participants. The transcripts are available in the supporting Information.

The third source of evidence is a post-activity questionnaire, administered anonymously using an online form. This was sent by email to all students in the cohort after the course deadline had finished, and 59 responses were received. The students were asked to answer the following questions (based upon those reported by Mac Raighne *et al.* (2015)) using a Likert scale:

1) Developing original question(s) improved my understanding of the topic(s)
2) Answering questions written by other students improved my understanding of those topics
3) Engaging in discussion on PeerWise (writing and reading comments) improved my understanding of the topics
4) I would use the questions available on PeerWise for revision purposes
5) I would like to see PeerWise used similarly again
6) I wrote my question(s) entirely from scratch and did not copy or paste them from elsewhere
7) The tutorial discussion improved my understanding of the topic(s)
8) Writing the question and answers was easy
9) A reputation score of 1000 was easily achieved
10) How much did your question change as a result of the tutorial?

The responses received are shown in Table 2.

As well as the Likert scale questions, three optional free-text questions were asked. These were as follows:
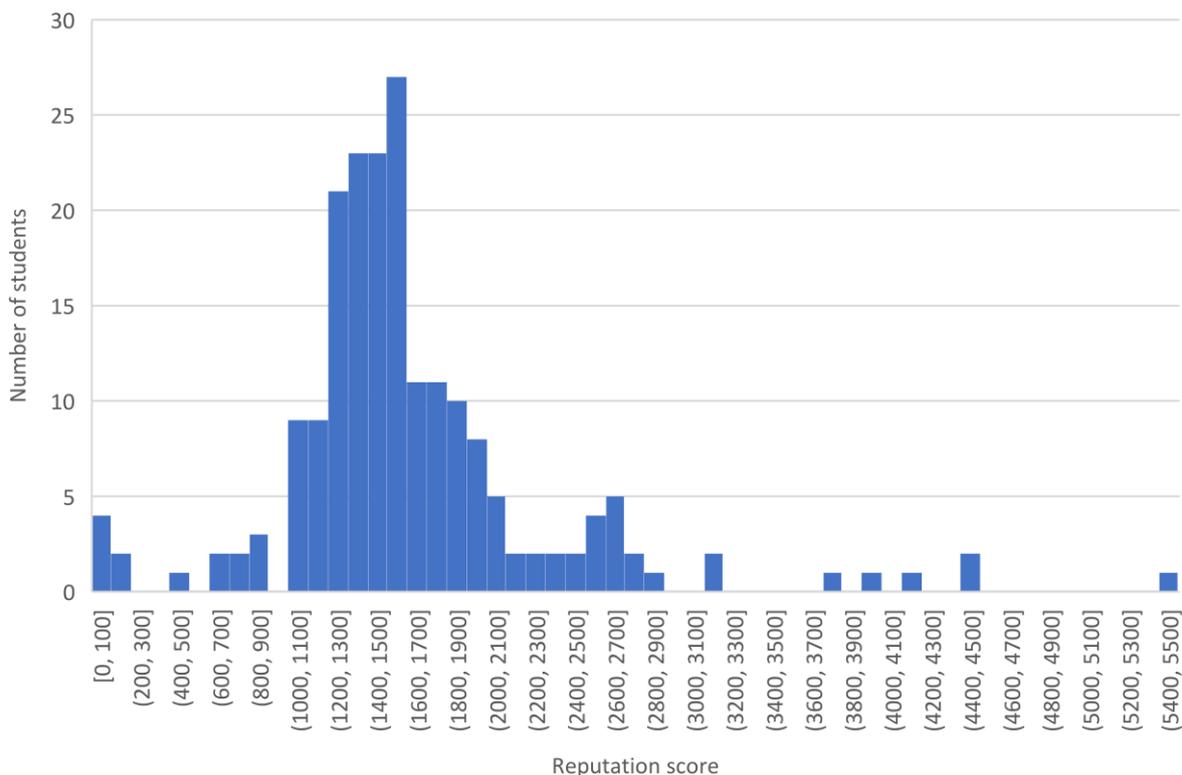
Figure 1 The distribution of reputation scores achieved by students.

11) What do you believe is the biggest benefit of using PeerWise? What aspects of using PeerWise did you find most useful/interesting/enjoyable?
12) What do you believe is the biggest problem with PeerWise? Can you recommend something that would make PeerWise more valuable or effective for learning in this class?
13) If you contributed more than the minimum requirement (either by developing more questions or by answering more questions than you were required to), why did you choose to do so?

Results and Discussion

Peerwise data

The answer to the first question appears to be that a target reputation score of 1000 is indeed a suitable target. Of the 215 students enrolled on the course, 198 registered with PeerWise, and 185 of these students achieved the target score of 1000. Of the 13 who did not, 5 displayed only a minimal level of engagement (i.e. did not upload their question, or did upload their question but did very little else), and the other 8 left their engagement until too close to the deadline. The lowest score to pass was 1009, and the highest scoring student scored 5408. 34 students scored more than 2000 points, and 151 scored between 1000 and 2000. The distribution of reputation scores is shown in Figure 1, and (as reported by others elsewhere (Ryan, 2013; Mac Raighne *et al.*, 2015; Fergus, 2019)) shows that generally participation was well beyond the bare minimum required for credit.

This study differs from many previous ones in that it asks students to achieve a particular score, rather than specifying that they should contribute a particular number of questions, answers and comments. This resulted in 406 questions in total (Table 3), 205% of the minimum total that would be expected from students simply posting one question each, though this was still the most common behaviour. Many more answers than questions were posted, presumably because answering questions is much less cognitively demanding: only 16% of questionnaire respondents agreed that writing a question was easy (Table 2), and this mirrors previous studies (Galloway & Burns, 2015).

	Questions authored	Questions answered	Questions correctly answered	Comments written	Ratings posted	Reputation score
Average per student	2.05	35.56	22.49	17.69	32.42	1630
Total	406	7041	4453	3503	6419	

Table 3 Aggregate data for student participation ($n = 198$)

Another finding from previous studies that is replicated here is that a small number of very active students make a disproportionately large contribution (Bottomley & Denny, 2011; Mac Raighne *et al.*, 2015). Thus in this study, 21 students (10.6% of the cohort) contributed between them 141 questions (34.7% of the total number of questions) with a maximum individual contribution of 14 questions. This contrasts with the 108 students (54.5.% of the cohort) who each only authored one question (26.6% of the total), presumably the question they took to the tutorial.

Tutorial recordings and questionnaire results

It was hoped that recording some of the tutorials would allow evidence for constructive evaluation to be obtained, and this proved to be the case. In the four tutorials recorded, which contained a total of 17 students, three students brought along questions that contained misconceptions. Student C submitted a question that assumed that two excited state atoms would combine to form an excited state diatomic molecule:

Tutor: So what is the diagram you've drawn for?

Student C: This is for two excited helium atoms. Becoming helium two, does that make sense?

Tutor: It kind of does. diagrammatically. But possibly not physically?

Student C: I can't really think of it in the physical sense, I just assumed it could happen because I read an article.

Student I was under the impression sp hybrids can only be formed along the z-axis:

Student I: sp hybrid orbitals are from p_z not from p_x

The tutor was able to explain that the choice of axis was arbitrary, though the z-axis was the conventional choice:

Tutor: the thing about the about x, y and z is that actually how you label your axes on a molecule is completely up to you. You know, you can orientate your axes x, y, and z however you like. I mean, by convention with z. You know, when we have the z, the z orbital with the d orbitals, that's the funny one with the donut, you know, we always call that the z one. But, but how we label things x, y, and z is up to us. So when you say, a p_y orbital can do something but a p_z orbital can't do that thing? That's not right. You know, because we could just orientate our axes the other way.

Student N was confused about the Bürgi-Dunitz angle, believing it to be solely due to the need to minimise electrostatic repulsion between reactants, and not appreciating that it arose from the overlap of orbitals:

Tutor: If it minimises electrostatic repulsion, why doesn't it attack from 90 degrees?

Student N: So I have to have both of them for it to be right?

Tutor: The point is that the electrostatic repulsion is part of the answer. But it's also that at 107, 108, 109 degrees, that's where the pi star orbital is - the orbital that the electrons are going into, and therefore that's, you know, how that reaction proceeds.

Comparison of the questions brought to these tutorials with those subsequently uploaded to PeerWise reveals that 9 of the 17 questions were modified before being uploaded, and this is supported by the answers to question 10 (Table 2), which indicate that overall more than half of questions were amended to some extent following the tutorial.

Furthermore, in discussing each others' questions, some of the students were also able to identify gaps in their knowledge. Student K wrote a question that talked about the HOMO of the $[BH_4]^-$ anion being a B-H σ-bonding orbital, which prompted student J to ask:

> …I was confused as to [option] B because I thought the like negative charge would be a lone pair. Why is it not?

And student B also wrote a question which mentioned the Bürgi-Dunitz angle. Following some discussion arising from it, other students in the tutorial said:

> Yeah, that makes a lot more sense. I think. I didn't know that [option] C was correct to be fair because that's just new information to me actually. (Student C)

> That clears up my confusion with the Bürgi-Dunitz angle. (Student D)

Students also reported choosing questions that targeted their zone of proximal development. Student M said:

> That's the reason why I chose this question, because I don't understand. And this is something I wanted to learn more about.

and student Q said:

> The first two answers are from some of the notes and I know I'm quite confused with.

As well as evidence from the recordings, the data in Table 2 show that the students self-reported that they learned through the activity, with 98% feeling that the process of writing questions improved their understanding of the topic, and 94% feeling that answering other students' questions on PeerWise did likewise. Several of the free text answers to question 11 also highlighted the learning benefit of writing a question. Examples include:

> The most useful and interesting aspect for me was writing questions. I also really liked the tutorial, as we had a proper discussion on our questions.

> Writing more questions allowed me to better engage with the material, which really helped me understand my weaknesses, while it was also a good revision.

> It was useful overall in developing my understanding by having to think of a question (instead of answer one) and there was a variety of questions and question types, which was useful.

> Creating your own question hugely helps to understand a topic - also having everyone else's question gives you ready-made revision material on lots of subtopics

The most common theme (6 comments) amongst the answers to question 12 about improving the activity was to do with its timing. PeerWise rewards prompt participation – the longer a question is in the system, the more points it can accrue for the author – and the timing of the tutorials presented a small hiccup. Students whose tutorials were later in week 4 had less time to engage with the system before the deadline at the end of week 5, making it more difficult to achieve the target score. Additionally, 30 questions were uploaded before the tutorial (despite students being asked not to), giving them extra time. However, 18 of these were subsequently deleted or amended. This 60% change rate contrasts with only 15% of the questions uploaded after the deadline subsequently being altered, implying that the quality-checking aspect of the tutorial is working.

The second most common theme in the answers to question 12 (4 comments from 59 questionnaire responses) was about the accuracy of questions, such as:

Lack of official moderation, sometimes an answer was wrong but very difficult to tell

However, previous studies into this have found that the incidence of wrong answers and explanations is low, (Bottomley & Denny, 2011; Galloway & Burns, 2015) and that these instances are normally picked up and flagged by participants (Mac Raighne *et al.*, 2015). One of the strengths of PeerWise is that it requires very little instructor input once underway; indeed, the absence of instructors allows the students to take ownership and create a safe space in which to make mistakes and to learn from them (Kay, Hardy & Galloway, 2020).

Conclusion

This proved a popular activity: 78% of respondents indicated that they would like to use PeerWise in a similar way again. Our implementation differs from many previous ones in that the tutorial is devoted to questions the students have already written rather than to scaffolding the activity, which takes place online and asynchronously. This seems to be a successful way to consolidate student learning on the chosen subject, by adding an extra opportunity for misconceptions to be revealed and discussed, and by increasing the quality and accuracy of the questions and the student confidence therein. The responses to questions 3 and 7 on the questionnaire indicate that students found the tutorial discussion much more valuable than any discussion that took place on PeerWise. Other studies have also reported that students find the online discussion less useful than writing and answering questions (Bottomley & Denny, 2011; Mac Raighne *et al.*, 2015). In the implementation reported herein, only a very small percentage of the comments on PeerWise actually discussed the chemistry in the question, and the overwhelming majority were bland praise (such as "Nice question!"); students seem to leave these as a means of easily increasing their PeerWise score, and next time the activity is used more emphasis will be placed on leaving constructive feedback, which students often overlook as a learning opportunity (Kay, Hardy & Galloway, 2018). The tutorial also seems to have largely prevented plagiarism, with only two questionnaire respondents raising concerns

about this, and only one admitting to having copied their question.

The whole activity has provided a bank of over 400 MCQs on the subject, some of which are of extremely high quality. In the questionnaire many students pointed out the utility of this as a potential revision resource. Interestingly, none of them pointed out that they might write more questions as a revision activity. As the creators of PeerWise have said, "Although they believe that writing questions helps them to learn, most students do not write questions unless they are compelled to" (Luxton-Reilly & Denny, 2010).

The only slight glitch related to the relative timings of the PeerWise activity and the tutorial. This can easily be remedied in future by simply not opening the course on PeerWise until all students have had their tutorial. However, the setting of a target score rather than a specific number of contributions worked, though instructors wishing to repeat this with different sized cohorts are advised to consult with PeerWise to identify a realistic score (Denny, no date).

Acknowledgments

References

Bates, S.P., Galloway, R.K., Riise, J. & Homer, D.(2014) *Assessing the quality of a student-generated question repository*, Physical Review Special Topics - Physics Education Research, 10(2), p. 020105. doi: 10.1103/PhysRevSTPER.10.020105.

Bottomley, S. & Denny, P. (2011) *A participatory learning approach to biochemistry using student authored and evaluated multiple-choice questions*, Biochemistry and Molecular Biology Education, 39(5), pp. 352–361. doi: 10.1002/bmb.20526.

Casey, M.M., Bates, S.P., Galloway, K.W., Galloway, R., Hardy, J. Kay, A.E., Kirsop, P. &

McQueen, H. (2014) *Scaffolding student engagement via online peer learning*, European Journal of Physics, 35(4). doi: 10.1088/0143-0807/35/4/045002.

Coppola, B.P. (2015) *Do Real Work, Not Homework*, in Garcia-Martinez, J. and Serrano-Torregrossa, E. (eds) Chemistry Education: Best Practices, Opportunities and Trends. Wiley-VCH Verlag GmbH & Co. KGaA, p. 236. doi: 10.1002/9783527679300.ch9.

Denny, P. (no date) *Scoring: for fun and extra credit!* Available at: https://peerwise.cs.auckland.ac.nz/docs/community/scoring_for_fun_and_extra_credit/.

Devon, J., Paterson, J.H., Moffat, D.C. & McCrae, J. (2011) *Evaluation of student engagement with peer feedback based on student-generated MCQs*, Innovations in Teaching and Learning in Information and Computer Sciences, 11(1), pp. 27–37. doi: 10.11120/ital.2012.11010027.

Fellenz, M.R. (2004) *Using assessment to support higher level learning: the multiple choice item development assignment*, Assessment & Evaluation in Higher Education, 29(6), pp. 703–719. doi: 10.1080/0260293042000227245.

Fergus, S. (2019) *Using PeerWise to support the transition to higher education*, in Seery, M. K. and McDonnell, C. (eds) Teaching chemistry in higher education : a festschrift in honour of Professor Tina Overton. Dublin: Creathach Press, pp. 181–193. Available at: https://overtonfestschrift.wordpress.com/.

Galloway, K.W. & Burns, S. (2015) *Doing it for themselves: Students creating a high quality peer-learning environment*, Chemistry Education Research and Practice, 16(1), pp. 82–92. doi: 10.1039/c4rp00209a.

Hardy, J., Bates, S.P., Casey, M.M., Galloway, K.W., Galloway, R.K. & Kay, A.E. (2014) *Student-Generated Content: Enhancing learning through sharing multiple-choice questions*, International Journal of Science Education, 36(13), pp. 2180–2194. doi: 10.1080/09500693.2014.916831.

Kay, A.E., Hardy, J. & Galloway, R.K. (2018) *Learning from peer feedback on student-generated multiple choice questions: Views of introductory physics students*, Physical Review Physics Education Research, 14(1), p. 10119. doi: 10.1103/PhysRevPhysEducRes.14.010119.

Kay, A.E., Hardy, J. & Galloway, R.K. (2020) *Student use of PeerWise: A multi-institutional, multidisciplinary evaluation*, British Journal of Educational Technology, 51(1), pp. 23–35. doi: 10.1111/bjet.12754.

King, A. (1990) *Enhancing Peer Interaction and Learning in the Classroom Through Reciprocal Questioning*, American Educational Research Journal, 27(4), p. 664. doi: 10.2307/1163105.

Krathwohl, D.R. (2002) *A Revision of Bloom's Taxonomy: An Overview*, Theory Into Practice, 41(4), pp. 212–218. doi: 10.1207/s15430421tip4104_2.

Luxton-Reilly, A. & Denny, P. (2010) *Constructive evaluation: A pedagogy of student-contributed assessment*, Computer Science Education, 20(2), pp. 145–167. doi: 10.1080/08993408.2010.486275.

Nicol, D. (2007) *E-assessment by design: using multiple-choice tests to good effect*, Journal of Further and Higher Education, 31(1), pp. 53–64. doi: 10.1080/03098770601167922.

PeerWise (no date, a) List of publications relating to PeerWise. Available at: https://peerwise.cs.auckland.ac.nz/docs/publications/.

PeerWise (no date, b). Available at: https://peerwise.cs.auckland.ac.nz/.

Race, P. (no date) *Designing Multiple-Choice Questions.* Available at: https://phil-race.co.uk/archive-of-downloads-from-previous-website/ (Accessed: 15 February 2021).

Mac Raighne, A., Casey, M.M., Howard, R. & Ryan, B. (2015) *Student Attitudes to an Online, Peer-instruction, Revision Aid in Science Education*, Journal of Perspectives in Applied Academic Practice, 3(1), pp. 49–60. doi: 10.14297/jpaap.v3i1.135.

Ryan, B.J. (2013) *Line up, line up: Using technology to align and enhance peer learning and assessment in a student centred foundation organic chemistry module*, Chemistry Education Research and Practice, 14(3), pp. 229–238. doi: 10.1039/c3rp20178c.

Talanquer, V. (2015) *Threshold concepts in chemistry: The critical role of implicit schemas*, Journal of Chemical Education, 92(1), pp. 3–9. doi: 10.1021/ed500679k.