

Multiple-choice questions in the Humanities: a case study of Peerwise in a first-year Popular Music course

Dr Adrian Renzo

Macquarie University

The web-based system Peerwise allows students to submit their own multiple-choice questions (MCQs) about course content, complete with distractors and an explanation of the 'correct' answer. Other students can then attempt their peers' questions and provide feedback on the quality of each question. To date, Peerwise has been used mostly in subjects where typical MCQs have a finite number of correct answers. This paper – a work-in-progress – suggests that Peerwise is potentially useful in units with a more 'discursive' orientation, such as MUS100, offered at Macquarie University, Australia. The system provides a good forum in which students can test each other on both lower-order and higher-order tasks. Future research will explore the efficacy of the tool across multiple iterations of the unit MUS100.

Keywords: multiple-choice questions, online assessment, peer review

Introduction

The web-based Peerwise system is potentially a useful complement to lectures and tutorials. It encourages students to review content and to engage with it in a more meaningful way than they might otherwise do individually. Peerwise requires students to submit their own multiple-choice questions (MCQs) about course content (complete with distractors and an explanation of the 'correct' answer). Other students can then attempt their peers' questions and provide feedback on the quality of each question. To date, Peerwise has most commonly been used in subjects where students need to systematically work through a set of problems, or where typical questions have a finite number of correct answers. For example, at Macquarie University, Peerwise is currently used by convenors in Business and Economics, Biology, and Software Engineering.

In contrast, this paper explores the possibilities of integrating Peerwise within the Humanities, where courses often have a more discursive orientation. For example, in the unit MUS100, students are expected to reflect on theoretical issues in popular music studies, demonstrate critical listening skills in relation to music production, and review the arguments of key theorists in the research field. At first glance, these aims strain against the MCQ format. The MCQ format seems better suited for topics in which a question has a clear, finite, and/or correct answer. In MUS100, the 'correct answers' to a single question may number in the dozens, and each 'answer' will be subject to the vagaries of individual students' comprehension skills and subjective interpretation. This paper argues that the discrepancies between typical uses of MCQs and the content of units such as MUS100 are not insurmountable. In fact, Peerwise provides a good forum in which students can test each other on both lower-order and higher-order tasks. In some cases, students begin to ask higher-order, 'meta' questions about a course, including why a lecturer takes a particular approach to teaching. The key point is that Peerwise remains useful even in situations where it is merely an early stepping-stone towards more substantial learning.

Lower-order and higher-order learning

Drawing on Marton and Säljö's seminal studies of the 1970s, Biggs (2011, p. 20) notes that students in higher education often adopt either a 'surface' or a 'deep' approach to learning. In the former, students may aim to memorise unrelated 'bits' of content provided by a lecturer, so as to amass a wall of knowledge which can then be used for summative assessment tasks such as examinations. The more academically oriented students tend to opt for 'higher-order' learning strategies. They attempt to connect new information with material they already know, bring questions with them to learning sessions, and apply ideas they have learned to new settings.

For some critics, the MCQ format tends to privilege lower-order learning. Karen Scouller (1998), for example, compared students' preparation for a multiple-choice exam and an essay-based exam. She found that an MCQ exam actually *discouraged* deep-learning: the people who adopted deep learning techniques paradoxically performed *less* well on the MCQ-based examination. However, that research assumed a conventional use of MCQ – one in which the lecturer was primarily responsible for generating the questions, and the students simply ticked (what they thought was) the correct answer. Several scholars have noted that the possibilities of MCQs

can be substantially expanded. Draper (2009) points out that there are several situations in which MCQs can be used to foster higher-order learning.

- Students can be asked to list reasons that each answer option is correct or incorrect, “rather than simply ticking one” (2009, p. 290).
- MCQs can be posed to students as the first step towards a broader discussion, rather than as an end unto themselves.
- Students can be asked to prepare their own MCQs.

Peerwise combines elements of all these strategies (Denny, Hamer & Luxton-Reilly, 2009). Students write their own question. Following the template that I provide, they also need to justify *why* each option is correct or incorrect. Finally, the system allows them to comment on each other’s questions, which is arguably where some of the ‘real’ learning happens.

How Peerwise ran in MUS100

In Session 1, 2014, I asked 156 first-year students to write one Peerwise question and answer five other questions, and to repeat the process several times during the semester. (In total, this involved writing four questions and answering twenty questions.) The assessment criteria were fourfold.

- Evidence that all aspects of the task have been completed.
- Evidence that the questions test people for deep learning rather than asking them to memorise isolated facts, figures, or quotes.
- Comprehensiveness of the explanation for each question.
- Level of prose, grammar, and syntax.

In practice, Criterion 1 simply measured the number of questions submitted. It is difficult to apply a quantitative value to students’ answers, because: (a) they may get the answer ‘wrong’ due to a misconception on the part of the question author, and (b) they may simply click on five boxes without reading the questions at all. Criterion 2 was added to discourage students from simply taking arbitrary quotes from a reading and asking: “who said this?” Criterion 3 was added in the hope that students would follow the template I established in my early questions.

What can Peerwise achieve for students in popular music studies?

There are several advantages of using Peerwise in a Humanities-based unit such as MUS100.

Peer review

Previous research has found that on systems such as Peerwise, higher-performing students tend to contribute more (Luxton-Reilly 2009). This was largely confirmed in the MUS100 iteration of Peerwise, in which students who earned Distinctions or higher in their written work reported being ‘addicted’ to Peerwise. (Nineteen students answered more than double the required number of questions; two students answered more than 300 questions.) On the one hand, such disproportionate contributions from students may be seen as giving us skewed results. However, as Luxton-Reilly notes (2009), the voluntary additional contributions from these students results in a higher overall level of feedback in any given cohort of students.

The element of peer review was also important for the convenor. Peerwise was a useful testing ground for questions which I would eventually place in an end-of-semester quiz. My experience here largely confirms Luxton-Reilly’s (2009) view that the disproportionate contributions from high-performing students raises the overall level of feedback. The feedback that I received on my own questions improved the quality of the questions which all the students eventually completed during in-class tests and quizzes. Here are two examples. In the first example, I asked students to apply what they had learned about music philosopher Theodor Adorno to a more recent example of popular music.

What would Adorno *most likely* say about Beyoncé’s music?

- A. After a week of assembling cars at a soon-to-be-closed car factory, workers need something that is entertaining without requiring any intellectual work. Beyoncé’s music is perfect for this.

- B. After a week surviving school and making cheeseburgers at McDonalds, teenagers need something to distract them from the homework that they should be doing. Beyoncé's music is perfect for this.
- C. Beyoncé's music is an example of standardised popular culture. Its purpose is to generate profits for wealthy industry executives while providing harmless, light entertainment for the masses.
- D. By releasing 'surprise' albums on iTunes, Beyoncé shows that she's aware of the challenges of marketing music in the digital era. In this way, she successfully generates profits for wealthy industry executives while providing harmless entertainment for the masses.

Explanation (excerpt):

B is incorrect because Adorno wasn't worried about kids doing their homework.

One student noted that my explanation did not adequately justify why option B would be 'wrong', and that the question might therefore be misleading. In another question, I asked students: 'What are the first three notes in the bass line of the song "Gun" by Chvrches?' One respondent wrote that this was confusing, because, strictly speaking, the first three notes are all the same, whereas I was obviously asking people to identify the first three *different* notes. In both of these examples, peer-review arguably strengthened the quality of questions that were later put to students in actual test situations, primarily by removing ambiguities which were not obvious to me (Malau-Aduli & Zimitat, 2012).

Gamification encourages competitiveness and improves motivation

At the end of the semester, I asked students to contribute one piece of advice for the subsequent cohort of students. One person's advice was: 'beat the lecturer at Peerwise'. This is a clear instance of gamification affecting students' responses to Peerwise. The system incorporates several elements from the gaming world: leaderboards give students an idea of how they are performing compared with other students, and participants earn 'badges' when they reach particular milestones, such as answering 10 consecutive questions correctly (Glover, 2013). The important thing here is that the rewards are aligned with other parts of the assessment structure, so that the person who (for instance) *reads* carefully achieves higher marks in Peerwise.

Peerwise can reinforce or cover material not emphasised in lectures

When more than 100 students are asked to submit several questions throughout the semester, there is obviously ample room for overlap (as, indeed, the topic cloud above demonstrates). When a point seemed to be emphasised in a reading or a lecture, many students ask (differently shaded) questions about that particular idea. In a cohort of this size, however, many students will voluntarily try to write more 'original' questions, by delving deeper into one of the prescribed readings, by capturing a seemingly peripheral point from a lecture, or by applying an idea from a lecture to one of their own examples. This proved to be one of the strengths of Peerwise. Rather than worrying about not adequately 'covering the content' in lectures, I could afford to make lectures more interactive, knowing that much of the material would be rehearsed and discussed online later in the week.

Peerwise can provide a meta-commentary on the learning process

One unintended outcome of using Peerwise in MUS100 was that students used the online setting to ask each other 'meta' questions about the content and the methods of content delivery in the unit. For example, I regularly briefed (Haynes, Haynes, Habeshaw, Gibbs & Habeshaw, 2012) the students on why I was running a particular segment of the lectures in a certain way. (For instance, this might involve drawing attention to aspects of my body language which should alert them to an important point worth noting, or explaining how to differentiate between essential and non-essential aspects of material displayed through PowerPoint.) Consequently, towards the end of the semester, some students contributed questions which required their peers to take stock of *how* they were learning, rather than simply asking for memorised answers to leading questions. Such questions might address why lectures were being used in the first place.

Potential pitfalls

One potential pitfall of using Peerwise is that students may use the system as a replacement for doing the required readings. Some students asked quite detailed questions about the prescribed readings, and it might be objected that the less academically-oriented students might try to grasp the main ideas of those readings by relying more on Peerwise. I would suggest that this risk can be mitigated by ensuring that Peerwise is appropriately aligned with other assessment tasks. In MUS100, for example, Peerwise was allocated only 5% of

the overall grade, whereas ‘tutorial worksheets’ (which included a series of questions about the required readings) were collectively allocated 35%.

The more significant pitfall lies in the tendency to over-assess students. For example, when the stakes are relatively low (the maximum mark for the assessment is 5%), I would suggest that it is cumbersome and unnecessary to include a criterion on students’ grammar. This is especially so when such aspects of students’ writing are already assessed in other tasks, and it would be especially important in any unit where a group of tutors were responsible for marking different groups of students.

Conclusion

The Peerwise activity used in MUS100 promises students that they will earn marks for submitting a certain number of questions, for writing a certain type of explanation, and so on. In practice, however, the real learning is arguably *incidental* to those assessment criteria. By being ‘prodded’ with the promise of marks, students willingly register on the Peerwise website and begin posting questions. Once they become involved in the process, at least some of them report becoming ‘addicted’ to the process and developing intrinsic motivation to participate. This raises interesting questions for future research. For example: is there any relationship between students’ Peerwise results and their overall grade? Do students who write many Peerwise questions improve their marks in other areas of the course, or does their Peerwise participation simply reflect their already-high engagement with the unit? While it is beyond the scope of this paper to provide a full analysis of the efficacy of Peerwise, this work-in-progress suggests that it may be productive to give Peerwise tasks some weight in the overall assessment schema – enough to serve as the initial ‘prod’ – and that it is counter-productive to reward such tasks with much more than 5% or 10% of the overall course mark.

Acknowledgements

This paper has been supported by funding through the Department of Media, Music, Communication and Cultural Studies (Teaching Index Funding), Faculty of Arts, Macquarie University.

References

- Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University*, (4th ed.). Berkshire: Open University Press.
- Denny, P., Hamer, J., & Luxton-Reilly, A. (2009). Students sharing and evaluating MCQs in a large first year Engineering course. *Australasian Association for Engineering Education Conference, 2009* (pp. 575-580). Retrieved from <http://aaee.com.au/conferences/AAEE2009/PDF/AUTHOR/AE090080.PDF>
- Draper, S. W. (2009). Catalytic assessment: understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology*, 40(2), 285-293.
- Glover, I. (2013). Play as you learn: Gamification as a technique for motivating learners. In J. Herrington, A. Couros, & V. Irvine (Eds), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2013* (pp. 1999-2008). Retrieved from <http://www.editlib.org/p/112246/>
- Haynes, A., Haynes, K., Habeshaw, S., Gibbs, G., & Habeshaw, T. (2012). *53 interesting things to do in your lectures*. Sydney: Allen & Unwin.
- Luxton-Reilly, A. (2009). A systematic review of tools that support peer assessment. *Computer Science Education*, 19(4), 209-232. Retrieved from https://www.cs.auckland.ac.nz/courses/compsci747s2c/lectures/paul/Luxton-Reilly_SystematicReview.pdf
- Malau-Aduli, B. S., & Zimitat, C. (2012). Peer review improves the quality of MCQ examinations. *Assessment and Evaluation in Higher Education*, 37(8), 919-931. Retrieved from <http://uq.edu.au/tediteach/assessment/docs/brief-26-may2012.pdf>
- Scouller, K. (1998). The influence of assessment method on students’ learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35, 453-472.

Please cite as: Renzo, A. (2014). Multiple-choice questions in the Humanities: a case study of Peerwise in a first-year Popular Music course. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology. Proceedings ascilite Dunedin 2014* (pp. 560-564).

Note: All published papers are refereed, having undergone a double-blind peer-review process.



The author(s) assign a Creative Commons by attribution 3.0 licence enabling others to distribute, remix, tweak, and build upon their work, even commercially, as long as credit is given to the author(s) for the original creation.