Dominik E. FROEHLICH, Daniela RING, Elisabeth JAHN

Multiple Choice Item Development Assignment in Lectures: A useful choice for multiple reasons

Abstract

In this workshop report, we present a newly designed multiple-choice item development assignment that draws from open educational resources. The major aim of the assignment was to enhance students' engagement and content-related learning as well as to help them to reflect on the very nature of multiple-choice questions. Thus a pool of multiple-choice questions was generated that was used as a practice and self-assessment opportunity for the students and from which several questions were picked for the final examination. First, the workflow of the multiple-choice item development assignment is described. Second, we discuss the advantages and disadvantages of this approach.

Keywords

multiple choice questions, lecture, open educational practices

1 Introduction

Assessing student's knowledge through multiple-choice questions is a widely spread practice in higher education (Nicol, 2007). Especially in lectures with a large number of participants, they are a common means of testing as the automatic evaluation enables quick and easy scoring in a more objective way than other types of assessment are able to (Douglas, Wilson, & Ennis, 2012).

Despite the prevalence of multiple-choice questions in many programs of higher education, they are a source of frustration for both students and teachers. On the one hand, multiple-choice questions often seem to test only superficial understanding of concepts, thus offering only a very shallow learning experience to the students (Scouller, 1998). On the other hand, competence can only be assessed in very limited ways (Nicol, 2007) while the development of good multiple-choice items is a very labor-intensive task (Haladyna, 2002).

In this workshop report, we present on a pedagogical intervention in the context of a lecture that aimed at (a) improving the content-related learning of students, (b) improving students' ability to tackle multiple-choice exams, (c) motivate students to participate in the course of the lecture, and (d) increase the overall quality of testing (Jahn, Ring, & Froehlich, 2019). For this, we use a multiple-choice item development assignment as presented by Fellenz (2004), but enriched with open educational resources and adapted to the specificities of the given lecture.

Fellenz (2004) introduced the notion of a multiple-choice item development assignment as a tool to engage students and improve their learning outcomes in higher education. This (graded) assignment—the submission of a given number of multiple-choice questions—was to be completed by students individually or in groups. The multiple-choice item development assignment was evaluated by the students as a positive learning experience, as it increases the time students spend studying for the course, helps them to learn more about how to prepare for multiple-choice exams, and reflect about what types of knowledge may be tested using multiple-choice questions. At the same time, Fellenz (2004) notes several limitations of the assignment, for example the high time requirement of both students and teacher.

We set out to further develop the multiple-choice item development assignment and implement it in a very different context: a public lecture without any requirement for attendance and potentially a large number of students. This difference is of importance, as it rules out the possibility to grade the assignment (which was a major motivator in the original study). Also, the problem identified by Fellenz (2004) concerning the amount of time needed to process the developed multiple-choice questions gets exacerbated given the potentially large numbers of students that attend a public lecture.

While the mentioned features make the implementation of the multiple-choice item development assignment more difficult than in the original scenario, there were also beneficial factors. Most importantly, we implemented the multiple-choice item development assignment in a course that was held the first time in a completely new curriculum (of educational science). This led to fewer constraints in the course design and gave somewhat more flexibility in terms of adjusting not only the assignment to the course, but also the course to the assignment. In addition, the newness of the lecture created a need to develop many multiple-choice questions—not only for examination purposes, but also for self-assessments and exercises.

2 The adapted multiple-choice item development assignment

The actual implementation was informed by two principles. First, the costs should be minimized and access for other teachers maximized by relying on open source software and other open educational resources. Specifically, a combination of a LimeSurvey questionnaire (Limesurvey GmbH, 2018), a R script (Froehlich, 2018a; R Core Team, 2014), and Moodle was used.

Second, the implementation needs to be sustainable. This means that the initial setup should be hassle-free, the scheme should be useable in a wide array of (educational) settings, and the process should not require a lot of manual interaction from the side

of the teacher. The workflow consisted of four major steps, which will be explained in detail in the following sub-sections.

2.1 Step 1

Students entered their questions in a LimeSurvey questionnaire (Limesurvey GmbH, 2018). In the questionnaire, the students identified themselves by the means of entering their student e-mail address. This step is optional in most settings, but since this was a public lecture, it was necessary to reward the students for their activity (see step 3 in this workflow). Then the students were able to submit one question and five answer alternatives. For each answer, they indicated whether the statement was true or false and added a verbal explanation or reference to justify the grading. This last step of entering verbal explanation of references to each answer is optional depending on the context. For us, the main motivation was to increase the effort it takes to submit questions to prevent quality problems. In addition, the verbal explanations or references aided the students' practice.

Several incentives were offered to encourage student's engagement in the development of multiple-choice items. First, all submitters of at least ten questions were given access to the pool of practice questions entered by the students. Second, we assured the students to consider the student-developed items when assembling the exam. Finally, as many as twenty percent of the exam questions were taken from the collection of items developed by students.

2.2 Step 2

After the deadline for entering questions passed, the LimeSurvey data was exported in .csv format. At this point in time, the teaching team did look at the developed multiple-choice questions to assess the overall quality of the items. During this (optional) process, we highlighted the questions that were of high quality for the potential inclusion at exams. Likewise, poorly written items were flagged for deletion.

After deleting a few questions, the data were processed by an open R script (Froehlich, 2018a) to transform the students' multiple-choice questions to Moodle's .gift format.

The students entered 583 questions, of which 363 (62%) questions were technically complete. Of the complete questions, 48 (13%) were flagged for deletion and 73 (20%) were flagged as superior (exam-ready) quality by two raters. The rest of 242 questions (67%) was judged to be adequate for the purpose of test preparation.

It makes sense to view these numbers in relation to the output of the teaching team, where two teaching assistants developed 55 questions throughout the course. This means that even if only the student questions of superior quality are counted, the item pool (per semester) is increased by 133%.

2.3 Step 3

The .gift file was imported in Moodle. Also, a list of students that submitted at least ten questions was extracted from the .csv file. These students were then assigned to a Moodle group that had access to all submitted questions through a Moodle quiz. Questions that were removed before the upload (see previous step) were anonymized and made visible to all students as a means of giving feedback on the multiple-choice question development itself. In addition, this list of “negatives” gave some information about what to expect—or rather what not to expect—for the examination (e.g., questions that were too detailed).

2.4 Step 4

After going through the questions of their peers, students were able to comment on questions or answers that they did not understand. This was the last line of quality management, where incorrect or strangely worded questions were flagged.

3 Reflection

The project has outcomes that may be classified in two large groups: outcomes for the students and outcomes for the teacher. In this section, we reflect about both types of outcomes.

3.1 Student outcomes

Concerning open learning and open education, the multiple-choice item development assignment can be seen as an interactive task that is based on open resources and technologies and creates practice material for other students. The most tangible benefit for the students is the practice opportunity generated by the 315 questions that were uploaded as practice questions. Since we report on the very first instance of a course, this is substantial opportunity for practice (as, based on our experience, new lectures often do not provide any practicing opportunities at all). While the sheer number of practice questions is likely to be beneficial to the students—for instance in terms of improving subject-related self-efficacy (Bandura, 1977) or getting a realistic self-evaluation before the examination—the heterogeneity of question developers may add an interesting additional dimension to this benefit. Since the students may look at different aspects of a text when reading it, the questions were posed from multiple viewpoints. This has the potential to make the evaluation better in terms of coverage. More research is needed to evaluate this hypothesis.

Content-related knowledge was (potentially) increased through engaging with the material in a more interactive fashion than usual for a lecture. This includes two major ways of learning. First, the reading process per se might be changed by imposing the “reading goal” of developing multiple-choice questions. Second, the feedback phase of the intervention (see Step 4 above) may lead to challenging others’ questions or being challenged based on a self-developed question. These instances of peer learning may help students to engage further with the material.

The intervention motivated students to do extra work without a direct influence on their grade. Their only reward was given through a different scoring system that allowed students to partake in enhanced exam preparation activities. As such, the intervention might also be interpreted in the light of the current discussion about the gamification of learning (Hanus & Fox, 2015). Indeed, many factors of a formal gamification system are present. However, we argue that the currency that is offered to the students—more subject-related learning and preparation opportunities—is more meaningful than points for rather short-lived scoreboards.

Public lectures usually provide no continuous assessment. The voluntary exercise encouraged continuous learning throughout the semester. By setting the deadline for submitting the items several weeks before the exam took place, students were encouraged to intensively deal with the relevant contents at an early stage. On a similar note, students mentioned that the exercise per se is engaging, as the development of questions is considered “a fun activity”. The list of questions that were deleted in the review phase was posted to the students. This allowed them to re-evaluate how multiple-choice questions should be posed and also what types of questions may be relevant in the course and which types of questions not.

Last, developing multiple-choice questions is a good opportunity to learn about and reflect about the very nature of such questions. This experience potentially makes it easier for students to interpret multiple-choice questions better and answer correctly (cf. Scouller & Prosser, 1994). This is an important competence especially for beginning students and may help this form of assessment to reach higher levels of acceptance among students. However, it may also further negative practices such as teaching (or learning) to the test (Popham, 2001; Volante, 2004).

3.2 Teaching outcomes

First, since many of the students’ multiple-choice questions were of excellent quality, the question pool was quite large, even though the course was offered the very first time in a completely new curriculum. This gave the instructors more freedom in assembling multiple-choice questions for the exam and thereby made the exam

more balanced. Through this extended pool of questions, stricter selection criteria were applied when reassessing the questions to pick for the final examination. From this perspective, it seems sensible to assume that the average quality of the multiple-choice questions posed did improve, too.

Second, sustainability and scalability were a big concern in this particular course. This is because the content of the course changes each semester. Subsequently, this means that also the texts relevant for the examination change every semester. Normally, this is not a good condition for multiple-choice questions, as the development and testing of multiple-choice questions is rather time consuming. Through engaging students in the multiple-choice developing process, the burden of developing multiple-choice items was shared with the students. This reduced the average time of the teaching team to develop items and hence made a multiple-choice exam feasible also in this setting.

Last, student engagement is an important result of the intervention. While this could be classified as a student-related outcome, we here focus on the opportunities heightened student engagement brings, especially in the context of a lecture without any requirements for attendance. Conversations with students revealed that those students that participated in the intervention engaged with the required reading not just before the exam but also during the semester. This plus in preparation on the side of the students may be used by allowing for a greater array of (more interactive) teaching techniques (cf. Froehlich, 2018b).

4 Conclusion

In sum, in this workshop report we reflected about two major components of a newly designed multiple-choice item development assignment that leverages open resources and open-source software. First, we have written about the assignment itself and reflected about the usefulness for both high education teachers and students. Based on the rather long and multifaceted list of advantages that were backed by

students and the teaching team, we are confident in recommending this type of exercise that creates actual practice material for other students. Second, we have shown how open resources and technologies have been used to leverage the multiple-choice item development assignment. Most importantly, the combination of standard open source survey software (Limesurvey GmbH, 2018) and a script that transfers the data in a format readable by a learning management system (Froehlich, 2018a) was an important ingredient in making this practice sustainable also in large classes.

5 References

- Bandura, A.** (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191.
- Douglas, M., Wilson, J., & Ennis, S.** (2012). Multiple-choice question tests: a convenient, flexible and effective learning tool? A case study. *Innovations in Education and Teaching International*, 49(2), 111–121.
- Fellenz, M.** (2004). Using Assessment to Support Higher Level Learning: the Multiple Choice Item Development Assignment. *Assessment & Evaluation in Higher Education*, 29(6), 703–718. <https://doi.org/10.1080/0260293042000227245>
- Froehlich, D. E.** (2018a). *csv2gift R Script*. Vienna. Retrieved from <https://github.com/dominikfroehlich/CSV2GIFT.R>
- Froehlich, D. E.** (2018b). Non-technological learning environments in a technological world: Flipping comes to the aid. *Journal of New Approaches in Educational Research*, 7(2), 88–92. <https://doi.org/10.7821/naer.2018.7.304>

- Jahn, E., Ring, D., & Froehlich, D. E.** (2019). Studentisch entwickelte Multiple Choice Fragen: A useful choice for multiple reasons [Multiple-choice items developed by students: A useful choice for multiple reasons]. Presented at the 2. Online-Tagung Hochschule digital.innovativ, Eisenstadt.
- Haladyna, T. M.** (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15(3), 309–334.
- Hanus, M. D., & Fox, J.** (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80, 152–161.
<https://doi.org/10.1016/j.compedu.2014.08.019>
- Limesurvey GmbH.** (2018). *LimeSurvey: An Open Source survey tool*. Hamburg: Limesurvey GmbH. Retrieved from <http://www.limesurvey.org>
- Nicol, D.** (2007). E-assessment by design: using multiple-choice tests to good effect. *Journal of Further and Higher Education*, 31(1), 53–64.
<https://doi.org/10.1080/03098770601167922>
- Popham, W. J.** (2001). Teaching to the Test? *Educational Leadership*, 58(6), 16–21.
- R Core Team.** (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org>

Scouller, K. M. (1998). The Influence of Assessment Method on Students' Learning Approaches: Multiple Choice Question Examination versus Assignment Essay. *Higher Education*, 35(4), 453–472.

Scouller, K. M., & Prosser, M. (1994). Students' experiences in studying for multiple choice question examinations. *Studies in Higher Education*, 19(3), 267–279.
<https://doi.org/10.1080/03075079412331381870>

Volante, L. (2004). Teaching to the Test: What Every Educator and Policy-Maker Should Know. *Canadian Journal of Educational Administration and Policy*.