

A Preliminary Study on the Learning Assessment in Massive Open Online Courses

Quan Yuan^(✉), Qin Gao, and Yue Chen

Department of Industrial Engineering,
Tsinghua University, Beijing 100084, China
{yuan-ql5, chenyl4}@mails.tsinghua.edu.cn,
gaoqin@tsinghua.edu.cn

Abstract. Massive Open Online Course (MOOC) is a new online education form. MOOC aims to provide the advance systematic educations to the public and share the access to the best high educations to Internet users. Although the MOOC platform contained many video lessons of high-quality courses from famous universities around the world, the assessment of students' learning, including testing methods, grading methods and feedback to student, was unsatisfactory according to XuetangX, an xMOOC websites leading by Tsinghua University in China. Setting effective and satisfactory assessment methods to test and grade students' learning performance in MOOC has significant values for all stakeholders including instructors, students and the MOOC platform. An interview study was conducted to understanding the current situation of assessments and the opinions towards different types of assessment methods from both instructors and students. We interviewed five teachers, eight course assistants of different categories of MOOCs in XuetangX, and six students from different MOOC platforms. Some conclusions and suggestions about the assessment on students' learning performance in different categories of MOOCs were drawn in the study. The findings in the study can be referred as guidelines for instructors to design great assessment methods in different MOOCs.

Keywords: MOOC · Learning assessment · Test method · Grading

1 Introduction

Massive Open Online Course (MOOC) is a new online education form [5]. MOOC aims to provide the advance systematic educations to the public and share the access to the best high educations to Internet users. Online users can take University courses on MOOC for free mainly by watching class videos, interacting with teachers and other students, doing homework, taking examinations and some other ways. MOOC is arousing global interest and leading to a new revolution in the education. Till today, hundreds of Universities have joined, thousands of courses have been provided, and millions of learners around the world have participated in varieties of MOOC platforms [4].

Currently MOOC is not very popular among people, and the high drop rate is a serious problem existing in the MOOC [10, 19]. The key characteristics of MOOC are openness, diversity, autonomy, connectedness and interactivity [7], and these arouse

many students joining in MOOC courses. The characteristics however bring three severe problems to MOOCs at the same time, particularly in student assessment [21]. Although the website contained many video lessons of high-quality courses from famous universities around the world, the assessment of students' learning, including testing methods, grading methods and feedback to student, was unsatisfactory according to XuetangX, an xMOOC [11] website lead by Tsinghua University in China. Because of the limitation of massive and open online education, a MOOC often had various students with different objectives and uneven knowledge level, and some of them had low level of knowledge and low learning skill but high study expectations [2]. It was hard for instructors to set appropriate courses and test for all students and for students to have motivations to learn the courses [18]. Consequently, teacher could not perfectly test students' ability and achievement as they did in real class.

Setting good assessment to test and grade students' learning performance in MOOC has significant values for stakeholders. For students, good assessments help them know the degree of knowledge acquisition, adjust their learning state, increase the sense of presence [1], improve learning motivations, and reduce the possibility of dropping out of the course [9]. For instructors, including teachers and course assistants, effective assessments can test students' ability, which can be a reference for instructors to adjust what they teach in time. For MOOC platforms, successful assessments can earn students' positive opinions of the course quality, and get the higher public acceptance to the course certification in MOOC platform. Thus, the ability to assessing learning performance of students is a necessary and important issue for the success of MOOCs.

This study attempted to: (1) investigate the current situation and requirements of assessment in MOOCs; (2) collect good example cases of assessment methods, summarize some suggestions about designing the test, grading and feedback; and (3) prepare for future work about proposing guidelines of how to design assessments for a MOOC. Firstly, we conducted interviews to instructors and students of different course types in XuetangX, to explore the applications of the different types of assessment methods to different courses, and the effects of those methods for evaluating students' performance. Secondly, we analyzed the data from interviews and concluded the effectiveness and satisfaction of the different assessment methods with varieties of courses in instructors' opinions.

2 Literature Review

2.1 Types of Tests in MOOCs

Assessment methods in MOOCs included objective questions, such as true-false, multiple-choice, filling in the blanks and matching, and subjective questions such as essay and product design. Group work and student engagement are separately special assessment method for offline course and online course.

Objective questions are the common assessment methods of students' performance both in traditional classes and in MOOCs. Objective questions test what extent the students remember and understand the knowledge [16]. It contains several common question types such as true-false, multiple-choice, filling in the blanks and matching,

suiting for examining closed-ended response and detailed knowledge [24]. Therefore, automatic judgment by computer can be conveniently done to objective questions, which are the benefit for MOOCs. Assessment consist only of objective questions for a MOOC has amounts of problems. Firstly, objective questions cannot test students about open-ended problems and their skills [12], and secondly, the test distinguishing degree of the questions was not good [3]. In addition, students can easily cheat in online objective question test. Besides objective questions, there should be other types of assessment methods used to judge students' performance in MOOCs.

Subjective questions are another common assessment methods. Those questions have no exact answer, usually asking students about their opinions to an open-ended item or designing a product such as painting, program and article. Subjective questions are used to test students' levels of understanding and using knowledge, and examine students' abilities of analyzing, summarizing and evaluating a problem and expressing their opinions [24]. Subjective questions are widely used in traditional education courses, but bring heavy checking burdens to instructors. As MOOCs are openness and generally arouse thousands of students to take the class, several instructors in one MOOC cannot check the subjective answers from such many students one by one.

Group work is a special type of subjective questions generally used in course assessment. Several students form a group to finish one assessment task, write one answer and hand it to instructors as their common assignment. Students in a group need high communications to finish one task. There were many restrictions in the group work. Firstly, students having different learning targets and attitudes are hard to come to an agreement about to what extent the assignment answer quality is. Secondly, a group member had dropped out from the course influence the process of the group work. Thirdly, students in a group can be difficult to gather together to discuss how to finish the work because of the time and place conflicts. In addition, a member's different consideration of the same work may be ignored by others in a group, letting that member giving up joining the group to work together with others [6]. Those problems stop the application of group work, particularly in MOOCs.

Student engagement is also a usual assessment in a course. In some traditional course, instructors frequently let students register their attendance at lessons, and take the attendance into consideration to grade students' final score. In MOOCs, the engagement of students is mostly measured by discussion. More discussions with higher quality bring better score to MOOC students. The engagement partly refers that how serious a student take part in the course, and is related to the students' performance, but some instructors thought the engagement makes no sense to students' knowledge level.

2.2 Grading Methods in MOOCs

The assignments are usually scored by instructors in traditional class. It is an effective grading method for courses which had small class size. Instructors have no enough time and energy to grading all assignments from a large number of students in MOOCs [22]. Even though they graded all the assignment by themselves, they had no time to give appropriate feedback to every student in time.

Auto-grading is used widely around MOOC platforms. The objective questions are effectively graded by auto-grading in XuetangX. But for subjective questions and other complex problems, it is often difficult to design machine grading systems which are accurate [20].

Peer review was a wide-used way of grading in MOOCs. It is a way to reduce instructors' checking burdens, and at the same time strengthen students' autonomy. It means a student's assignment is grading by his or her classmates. It has been used in Coursera, one of the most popular MOOC platforms. The knowledge gap among students in MOOCs may cause the problem of reliability and validity in peer review [13].

2.3 The Criteria of a Valid Assessment

The validity of an assessment includes both the effectiveness and satisfaction to the assessment, and the two items can come from both instructors' expert evaluation and students' judgment.

The effectiveness of an assessment includes [6, 15, 23]:

- if the assessment can examine the teaching knowledge of the course
- if the assessment can examine students' abilities of integrating knowledge and applying knowledge to analyze and solve a problem
- if the assessment can reflect the learning improvement of students

The satisfaction of an assessment is related to the following factors [13, 25]:

- if the assessment can motivate students for harder and better studying
- if the instructors offer helpful assessment feedback in time in studying process
- if the assessment is clear and fair

3 Interview

The objectives of interview research were to understand the overall assessment status and requirements of the MOOCs, and concluded some useful design guidelines for better assessment of MOOCs.

There are some factors that are related to the effectiveness and satisfaction of an assessment, and need to be figured out in the interview. Xia and Jiang [24] stated that education evaluation should consist of cognition, skills and affection, where the general method to measure cognition and skills was test. Firstly, at the student angle, the motivation to take a MOOC class will influence the completion of the course [17], and the foundation knowledge level of students about a specific course affects learning outcomes and further affects assessing and grading. Secondly, the assessment of a course and the feedback of the assessment can change students' learning motivation [9, 14]. Moreover, the types of course and teaching objectives influence the design of assessment methods and contents, and then affect the validity of the assessment.

3.1 Interview Process

Course Classification. Different types of assessment method fit for different courses, and similar types of assessment method usually fit for same categories of courses. As the instructors often used different assessment methods according to different types of courses, the first step preparing for an all-sided interview was classifying courses into different categories at XuetangX.com. Referring to national standard classification of disciplines, and considering the current categories of XuetangX at the same time, we divided MOOCs into eight categories: (1) science, (2) medical science, (3) computer science of engineering, (4) other engineering, (5) business and economics, (6) art and design, (7) language, (8) other liberal arts.

Participants. 13 instructors and 6 students of MOOCs in XuetangX.com were interviewed in the research. The instructors included five teachers who set up at least one courses of different categories on XuetangX, and eight assistants of MOOCs in eight different categories. Table 1 showed the interviewing instructors' courses and categories information. C++ Program Design and English Listening & Speaking were two courses of which we interviewed both teacher and assistant. As the table showed, we encoded the instructors with different categories 1–8 and whether he or she was I. teacher or II. assistant. For example, 3-I referred a teacher for computer science. The students were all XuetangX users from different place in China. All of them had study or were studying one or more MOOCs belonging to different categories. Some of students had experiences on studying on other MOOC platforms such as Coursera and edX.

Table 1. The courses of interviewing instructors

Category	I. Teachers	II. Assistants
1. Science	<i>Modern Biology</i>	<i>Linear Algebra</i>
2. Medical science	<i>Medical Parasitology</i>	<i>Introduction to Medical</i>
3. Computer science	<i>C ++ Program Design</i>	<i>C ++ Program Design</i>
4. Other engineering	None	<i>Circuit Theory</i>
5. Business and economics	None	<i>Financial Valuation Analysis</i>
6. Art and design	None	<i>Art of Eternity</i>
7. Language	<i>English Listening& Speaking</i>	<i>English Listening & Speaking</i>
8. Other liberal art course	<i>Basic Principle of Marxism</i>	<i>History As A Mirror</i>

Interview Questions. We interviewed teachers, assistants and students with different open questions. The teachers were interviewed about their teaching objectives, course content, student information, current assessment methods and their opinions about all kinds of assessment in their MOOCs. We asked assistants about the course content, assistant tasks, details about students, current assessment methods, opinions about all kinds of assessment and grading methods. Students expressed their objectives they wished to get from the MOOC, the current situations about course teaching and assessment of different MOOCs, the opinions whether a specific assessment methods fit for those MOOCs, and the suggestions that they thought the course assessment should be.

Procedure. All interviews for instructors were conducted face to face. All interviews for students were conducted through online phones. The interview time for each participant was about 20 min.

4 Results

4.1 Instructors

Current Situation Analysis. Table 2 showed the teaching objective, teaching content and student information in different MOOCs according to instructors. The left column of the table contained several descriptions related to the course, and other columns referred different courses and their true and false about each description. The course is coded as we mentioned above, 1–8 for categories and I/II for the identities of participants. Y referred that the description was true in the course, and N referred that the description was false in the course, and the white piece referred that the description was not mentioned by instructors.

Table 2. Current situations in different MOOCs

Descriptions	1-I	1-II	2-I	2-II	3-I	4-II	5-II	6-II	7-I	8-I	8-II
Low teaching objectives	Y	Y		Y	Y	Y	Y		Y		Y
Simple study content	Y		N		N				Y		
Simple assesment	Y		Y		Y			Y	Y	Y	Y
No course improvement	Y	Y		Y	Y				Y		
Big gap of students'basic level	Y	Y	Y	Y	Y	Y	Y		Y		Y

We got the following course features:

- The MOOCs were simpler than off-line classes in both teaching contents and ways to assessment, and were not improved during the semesters changing
- Instructors generally knew less about students' learning status, including large amounts of students and the big gap of students' basic knowledge level
- Students were not active during course
- The MOOCs were regarded as a resource but not an online class

Current Assessment Method. Table 3 showed current assessment methods in different MOOCs. The structure and coding of the table was similar with Table 2, with the descriptions, courses and true/false information from instructors.

We got the following assessment features:

- All of the courses whose instructors were interviewed used only objective questions to test students' performance whatever the course categories were, and the objective questions used included true-false, multiple-choice and filling in the blanks, with

Table 3. Current assessment in different MOOCs

Descriptions	1-I	1-II	2-I	2-II	3-I	4-II	5-II	6-II	7-I	8-I	8-II
Homework	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Final exam	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
Discussion bonus	N	N	Y	N	N	Y	Y	N	N	Y	N

automatic check and without feedback from the instructors to students. No subjective questions be applied by the courses

- Half of instructors considered that students' bonus on discussion were helpful to measure students' active level and speech quality, but were not applied by instructors because of the suggestion from XuetangX that the scores were too subjective to be used in the MOOCs. There were three courses had cut the discussion bonus they had set up before. Instructors expressed their dissatisfaction about the decision made by XuetangX

Opinions About Different Assessment. We introduced opinions to different assessment method separately.

Opinions about Objective Questions.

- Majority of the instructors agreed that the most important role the objective questions played was examining whether a student had watched the uploaded teaching video in a MOOC
- Most art and design and liberal art subjects cannot validly test students' performance only with the objective questions, while science and engineering courses' instructors thought the objective questions were enough for them to test students' actual learning level, according to the instructors of *Linear Algebra*, *Circuit Theory*, *Financial Valuation Analysis*, *Modern Biology* and *Introduction to Medical*
- The objective questions mostly fit for students' basic knowledge level, according to the instructors of *English Listening and Speaking* and *Medical Parasitology*
- Students were able to cheat easily when answering the objective questions
- All the assistants thought helpful feedback in questions could reduce the workload of instructors from answer questions, and improve students' learning outcomes

Opinions about Subjective Questions.

- All the instructors affirmed the instructive and positive effect of subjective questions.
- The most important role the subjective questions played was examining a students' true learning status, skills and abilities the course taught in most liberal art subjects in a MOOC, such as *Basic Principle of Marxism*, *History As A Mirror* and *Art of Eternity*
- In most science and engineering courses, such as programming and proof problems, subjective questions can test the knowledge and skills, but they were too hard for all the students to complete it as an assignment

- XuetaangX thought the discussion bonus was violated the object evaluation of the students' learning status, and did not support it
- The subjective questions would bring heavy burden to instructors as they always need to grade all assignments, and the number of students in one MOOC was always a large amount

Opinions about Discussion Bonus.

- Majority of the instructors considered that the most important roles the discussion bonus played were examining whether a student made an effort to learning the course, and provided an active atmosphere in MOOCs
- Most art and design, liberal art and biology courses need test students' performance only with the discussion, while the engineering courses' instructors thought the discussion cannot test students' actual learning level
- XuetaangX thought the discussion bonus was violated the object evaluation of the students' learning status, and did not support it. There actually existed the course that needed students to discuss. The discussion bonus should not be completely eradicated
- The discussion bonus was hard for instructors to grading. MOOC platforms should add a function supporting to collect students' discussions for instructors

Opinions about Group Work and Peer Review. Group work was anxiously needed by the courses that focused on opinion changing, communication and discussion, including language, art and design and social sciences subjects. But the group work was the advantages brought by small size class teaching in traditional education. Instructors were worried about the difficulties during the implementation and application of group work in large-scale teaching such as MOOCs. With less instructors' leading and more students and groups, group work might cause little attention to one specific student, which did worse than normal assessment. And group work would bring a large heavy burden to instructors, especially to assistants. Moreover, there were no successful precedent of group work used in MOOCs, so instructors were dare to use it.

Half of the instructors gave positive attitudes towards the peer review's effects on solving the grade process to judge subjective questions, and reducing the burden of assistant to check all the answers students submitted. The implementation of the peer review was ever hard. The instructors thought there must exists students that had no enough knowledge to assess other students' assignment, and students refused to grade other works and even give the score at random, which let the score finally need to be graded again by instructors. There were successful cases in *English Listening and Speaking* and *Art of Eternity* for peer review. Detailed grading criterions and good punitive measures should exist for peer review. The punishment to a student who scored other works at random should be established. An assignment should be distributed to several students, verifying there were at least a serious score given to the assignment.

4.2 Students

Learning Objectives. There were two main ideas about why the six students began studying a MOOC: (1) They were interested in the topics of the course; (2) They wanted to get useful knowledge or skills in the MOOC so that they can use what they had learned in future. The students who had different objectives chose different categories of MOOCs and had large different expectations for what they can get from the course. Those who studied for interest preferred (a) art and design and (b) liberal art courses, such as *News Photography* and *Introduction to Psychology*. Engineering, business and economics and language categories were more likely chosen for studying useful knowledge, for example, programming courses and English courses.

Opinions about Current Assessment. There almost existed no subjective questions in all courses the six students studied in XuetangX. For those studying for useful knowledge, they thought that objective questions were simple in most cases, and could not test their true grade and brought no positive effects on promoting their motivations to study. But in some science and engineering course, objective assignments could be hard to do. Some particular cases were mentioned that subjective questions were used. The good examples are photos in *News Photography* and story discussing in *Introduction to Psychology*. The good subjective questions aroused students' interest and sense of participation during studying. One of them suggested that the subjective questions and answers could be provided as materials but not assignments, so that students who need more exercises used it, and teachers had no need to grade the questions. For liberal art courses, course forums and online instant chat services were considered to be good places to discuss the subjective topics.

Opinions about Other Assessment. All of students had positive attitudes towards group work and peer review. They thought both could increase the communications with students, and improve the sense of participation and motivations to study. Appropriate laws should be set at the same time to better assessment. Valid test and acceptable work burden could be achieved for both instructors and students. One students said it would be better to test different questions with different difficulty for different students, and students could choose assessment according to their own needs.

5 Conclusion

Some findings about the assessment on students' learning performance in MOOCs were prompted in the study:

Firstly, there actually existed some types of courses as follows whose assessment needs can be met to test students' performance:

- The courses that primary problems are solved by analytic methods and having only one correct answer, including (a) parts of science, (b) engineering and economics subjects
- The courses focusing on examining fixed knowledge point in details which need to remember by students, including (c) some science and (d) medical science subjects

The subjective questions can also be used in MOOCs we referred above to get better evaluation.

Secondly, the subjective questions were necessary in the following courses:

- The courses need students' expression about their opinions to topics and submitting their products, including (a) social sciences, (b) art and design subjects
- The courses requiring students to improve themselves through communication with others or group work to complete a task, including (c) language and (d) some medical science subjects

Subjective questions in most courses were difficult for most students to finish. Instructors can offer the subjective questions to students as reference materials. The completeness of those questions were not the criterions to students' final scores.

Thirdly, Students' discussion played a key role to activating course atmosphere and promoting students' expression and communication. It was able to work on most of liberal arts courses. Course forum and online instant chat services were the way to discuss.

Fourthly, feedback for some difficult test and discussions was necessary, including different arguments about a topic on liberal art courses and common mistakes on science and engineering courses.

Fifthly, group work and peer review were useful for most categories of courses but hard to use. Peer review could strengthen students' autonomy. Grouping and burden problem should be solved in future. Detailed grading criterions and good punitive measures were the current ways to better peer review.

In addition, designing different degree of test difficulties might be the solution for the diversities of students' learning objectives and knowledge levels.

The findings in the study were guidelines on how to designing better assessment methods of MOOCs. More details about building valid assessments in a specific MOOC need to be intensively studied according to the interview. A handbook including guidelines and attentions about how to design the assessment in MOOCs should be proposed and verified in future work.

Acknowledgement. This study was supported by the Online Education Research Foundation of the Online Education Research Center, Ministry of Education, P.R. China (Project No. 2016ZD103) and the National Natural Science Foundation of China (Project No. 71401087).

References

1. Anderson, T., Dron, J.: Three generations of distance education pedagogy. *Int. Rev. Res. Open Distrib. Learn.* **12**(3), 80–97 (2010)
2. Belanger, Y., Thornton, J.: Bioelectricity: A quantitative approach Duke University's first MOOC (2013)
3. Chen, X., Dai, S., Zhao, X., Bai, Z., Liu, Q., Zhi, S.: Efficiency analysis of different question types in a medical imaging examination (in Chinese). *Chin. J. Med. Educ. Res.* **11**(11), 1160–1163 (2012)

4. Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., Emanuel, E.J.: The MOOC phenomenon: who takes massive open online courses and why? (2013). SSRN 2350964
5. Cormier, D.: The CCK08 MOOC–Connectivism course, 1/4 way (2008)
6. Cross, S.: Evaluation of the OLDS MOOC curriculum design course: participant perspectives, expectations and experiences (2013). <http://oro.open.ac.uk/37836/>
7. Downes, S.: Connectivism: A theory of personal learning (2008). <http://www.slide-share.net/Downes/connectivism-a-theory-of-personal-learning>. Accessed on July 14, 2014
8. Harmon, O.R., Lambrinos, J.: Are online exams an invitation to cheat? *J. Econ. Educ.* **39**(2), 116–125 (2008)
9. Hsia, L.H., Huang, I., Hwang, G.J.: Effects of different online peer-feedback approaches on students' performance skills, motivation and self-efficacy in a dance course. *Comput. Educ.* **96**, 55–71 (2016)
10. Kolowich, S.: Coursera takes a nuanced view of MOOC dropout rates. *The chronicle of higher education* (2013)
11. Kop, R.: The challenges to connectivist learning on open online networks: learning experiences during a massive open online course. *Int. Rev. Res. Open Distrib. Learn.* **12**(3), 19–38 (2011)
12. Krathwohl, D.R.: A revision of Bloom's taxonomy: an overview. *Theory Pract.* **41**(4), 212–218 (2002)
13. Luo, H., Robinson, A.C., Park, J.-Y.: Peer grading in a MOOC: reliability, validity, and perceived effects. *J. Asynchronous Learn. Netw.* **18**(2), n2 (2014)
14. Maier, U., Wolf, N., Randler, C.: Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Comput. Educ.* **95**, 85–98 (2016)
15. Manalo, J.M.A.: An Evaluation of Participants' Levels of Satisfaction and Perceived Learning Regarding the MOOC in@ RAL Platform. *Malays. J. Distance Educ.* **16**(1), 101–121 (2014)
16. Miranda, S., Mangione, G.R., Orciuoli, F., Gaeta, M., Loia, V.: Automatic generation of assessment objects and remedial works for MOOCs. In: 2013 International Conference on Information Technology Based Higher Education and Training (ITHET), pp. 1–8. IEEE (2013). http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6671018
17. Nikola, S.: Effects of motivation on performance of students in MOOC. In: International Scientific Conference of IT and Business-Related Research-SINTEZA, pp. 418–422. Singidunum University (2014)
18. Onah, D.F., Sinclair, J., Boyatt, R.: Dropout rates of massive open online courses: behavioural patterns. *Proceedings on EDULEARN14*, pp. 5825–5834 (2014)
19. Parr, C.: MOOC completion rates 'below 7%'. *Times Higher Education*, 9 (2013)
20. Readers, H.: Professionals against machine scoring of student essays in high-stakes assessment. *Human Readers* (2013)
21. Sandeen, C.: Assessment's place in the new MOOC world. *Res. Pract. Assess.* **8**, 5–12 (2013)
22. Shah, N.B., Bradley, J., Balakrishnan, S., Parekh, A., Ramchandran, K., Wainwright, M.J.: Some scaling laws for MOOC assessments. In: *KDD Workshop on Data Mining for Educational Assessment and Feedback (ASSESS 2014)* (2014)
23. Suskie, L.: *Assessing student learning: A common sense guide*. Wiley, Hoboken (2010)
24. Xia, H., Jiang, S.: *Educational Psychology*. Tsinghua University Press (2015). (in Chinese)
25. Yousef, A.M.F., Chatti, M.A., Schroeder, U., Wosnitza, M.: What drives a successful MOOC? an empirical examination of criteria to assure design quality of MOOCs. In: 2014 IEEE 14th International Conference on Advanced Learning Technologies (ICALT), pp. 44–48. IEEE (2014)