

An Empirical Study on the Effects of Pedagogical Intervention on Improving the Quality of Peer Assessment in Massive Open Online Courses

大規模オープンオンライン講義における教育的介入がピア評価の質的改善に及ぼす効果に関する実証的研究

Nikan Sadehvandi

1. Statement of the Problem

Massive Open Online Course (MOOC) has grown substantially in recognition over the past few years. MOOC provides higher education leaders and institutions with the great impetus to share content and adapt online learning to address the needs of non-/traditional students worldwide. While MOOC offers opportunities for liberated and self-directed learning, there are still challenges with its pedagogy that need to be addressed for MOOC to become a more sustainable educational model. As an integral part of MOOC instructional design, one area that needs more empirical research is peer assessment. This is mainly due to the fact that the logistics of peer assessment in MOOC are different from conventional practices. In turn, this type of peer assessment could be readily affected by assorted factors such as learners' assessment bias, lack of experience or expertise in assessment, different educational background, and lack of motivation to do a quality assessment.

2. Purpose of the Study and Research Questions

In line with what was stated above, the study is conducted as an attempt to contribute to the understanding of this method of assessment and possible measures which could be taken to minimize, if not completely eliminate, the impact of such adverse factors.

This study addresses the following research questions:

1. What are MOOC learners' perceived issues with peer assessment? What do students think of the necessary conditions for optimizing peer grading accuracy and fairness in MOOCs?
2. How does additional support on MOOC peer grading affect the quality of grades in terms of accuracy? How does learner's factors affect the quality of peer assessment in terms of accuracy? How does additional support affect MOOC learners' perceived fairness of and satisfaction with their grades?
3. Does incentivizing peer graders have any effects on improving the quality of MOOC peer assessment in terms of accuracy? How incentivizing peer graders affect the time spent on grading peers? How does incentivizing peer graders affect MOOC learners' perceived fairness of and satisfaction with their grades?

3. Peer Assessment in MOOC

A review of the related literature on MOOC with a particular focus on peer assessment was initially conducted in this study. Central to the literature review was the inquiry into how peer assessment is currently being approached in MOOC and what needs to be addressed to improve its quality. The review of the literature unraveled benefits associated with peer assessment, challenges in its implementation, and some the recent efforts conducted to address these challenges especially in a MOOC. The concern that assessment of learning outcome in MOOC does not sufficiently capture students' performance and competency was raised and several trends on assessment were introduced and published studies relevant to each trend was covered. Furthermore, although the commonly used software programs such as an automated essay scoring tool have shown potency, they are still limited in terms of gauging learners' analytical reasoning and critical thinking. In fact, evaluating those essays that involves the use of such skills can be quite subjective, topic-dependent and calls for the same set of skills to assess the essay on a deeper level than merely attending the syntactic structure or relying on word count.

As for peer assessment, the literature review provided evidence on the need for more research on the effectiveness of this method in MOOC and the necessary conditions for its outcome to be valid. Two reasons were given on why the research on MOOC peer assessment has remained underrepresented: (1) the fact that MOOC history is brief and the evidence on the effectiveness of carefully designed experiments is scarce, (2) there has been a great urge to tackle the issue of peer assessment quality from a computer programming rather than an education paradigm.

However, such projects are often budget consuming, require a good amount of technology start-ups, and a great deal of time and expertise on the part of the MOOC design team. Moreover, the potential application of research findings from the pre-MOOC peer assessment era to feed into and base the backbone of designing, implementing and testing peer assessment in MOOC should not be overlooked. Training and support as one of the variables, which is deemed to enhance the accuracy of peer assessment results, have not been fully embraced in the pedagogical design of MOOC. Therefore, research on adopting pedagogical interventions to scaffold and support learners during the peer assessment activity and to study the possible impacts of such pedagogical approaches on reliability and validity of peer assessment results is required. Finally, the importance of motivating learners for conducting a quality peer assessment and taking to account their perception of peer assessment are also discussed.

4. MOOC Learners' Perception about Peer Assessment

This research also delved into learners' perception and views about peer assessment conducted in KyotoUx:001x as a case study. The purpose of this case study was to grasp learners' challenges and perceived issues and take away lessons and suggestions that could be applicable not only for improving peer assessment guidelines and implementation in forthcoming iteration of the same course, but also to give more universal implications that

could be implemented in other courses. To this end, learner's' discussion forum posts associated with peer assessment and their suggestions on the course ending survey was used as the main corpus of data in the analysis. The result of a qualitative analysis of students' data suggested nine key themes in students' posts: (1) improving grading distribution on the rubric, (2) addressing problems with the use of technology in the course, (3) better implementation of prize, (4) improving the grading rubric itself, (5) mandating written feedback on grading, (6) making known the identity of peer assessors, (7) monitoring the grades given by the staff, (8) inclusion of instruction and training on peer assessment, and (9) Penalizing those who grade poorly. Some of these suggestions were in line with students' responses on the course ending survey: (1) Rethinking the grade distribution on the rubric, (2) Training and instruction on the use of rubric, (3) Encouraging fair peer assessment, (4) Requiring written feedback. The case study also puts forth several implications for improving peer assessment in future MOOCs. From the authors' perspective some of the issues pinpointed by the students could be course-specific, some others exist across MOOCs. Thus, addressing these issues might have implications for designing and improving peer assessment in other MOOCs. Based on the literature review and the case study described above, the remainder of the research empirically examined the effects of two interventions on raising the quality of peer assessment in MOOC.

5. The Effects of Instruction and Training on MOOC Peer Assessment Quality

The first intervention was conducted to study the effects of different levels of learners' involvement with a support mechanism on peer assessment designed to raise the accuracy of peer grading and its effects on learners' perception of fairness and satisfaction with their grades. Additionally, the roles of two other learner attributes: prior experience with MOOC peer assessment and the time spent on the peer grading were also investigated. The study was conducted using peer grading data on a homework assignment in KyotoUx:002x. The support mechanism consisted of a training step and a video instruction, which were imbedded into the courseware and activated along with the peer assessment activity. The training step provided a chance for learners to practice with the course peer grading rubric before proceeding with the peer assessment activity. The video instruction also gave a detailed description of the purpose of peer assessment, a step-by-step guide on how to use the training step, and an extensive elaboration on the course peer grading rubric and its scales.

There were three experimental conditions in the study: (1) Full-treatment condition that consisted of peer graders who took the training step and watched the video instruction, (2) Partial-treatment condition that was made up of peer graders who only watched the video instruction, and (3) Control condition that included peer graders who neither took the training step nor watch the video instruction. However, each submission was graded by at least two peer graders who might have correspond to one of the three experimental conditions.

Moreover, because the distribution of submissions to peer graders on edX was done by the default peer review system, prior selection of peer graders into different experimental conditions was not an option in this study. As an alternative, it was set to find what experimental condition the two peer graders of the same submission belonged to. This

observation yielded five patterns which was called “Case Types” and the analysis was conducted comparing these Case Types.

The results of the analyses indicated that peer grader agreement was higher in those Case Types which had higher levels of involvement with the support mechanism. Also, there were significantly higher correlation indices between peer graders more involved with the support mechanism and the teaching assistant with respect to their grades on the same submission. The time spent on grading and prior assessment experience did not turn out to be major influencing factors on students’ accuracy of grading in this study. Also, students were mostly satisfied with their grades and perceived them as fair.

6. Incentivizing Peer Assessment Quality in MOOC

The second intervention had to do with an incentivizing mechanism, dubbed virtual prize group (VPG). It aimed to improve the quality of peer assessment in MOOC through taking a different approach to utilizing the otherwise equivocal anonymity feature of MOOC combined with *best performance award* offered to a group of peer graders, the mechanism had two purposes.

Firstly, it attempted to improve the accuracy of peer awarded grades on a homework assignment in KyotoUx:002x MOOC. Secondly, it sought to encourage students to spend more time on the peer grading activity. The core attribute of the VPG mechanism was to allocate a prize on a group-based format as opposed to awarding it to individual peer graders while keeping the identity of the groups anonymous. Learners who submitted their work were informed prior to the advent of peer assessment that they each belonged to a randomly predefined group and only one of these groups would receive the prize as a group-based accomplishment. However, the identity of the group members was neither revealed to the rest of the group nor other learners in the other rival groups.

The prize was designated for the best performance on homework assignment. Ultimately, only one group whose sum of members’ grades on this homework was higher than the other groups was selected as winner and awarded the prize at the end of the peer assessment process. Simultaneously with the launch of the homework on the courseware, students were informed about the VPG mechanism through the instruction video for the homework, course update, newsletters, and the discussion forum specified thread.

The results indicated that although the Virtual Prize Group (VPG) mechanism did not have any effect on improving peer grading agreement, it encouraged peer graders to spend more time on the grading activity. Also, most learners indicated their overall satisfaction with their grades and perceived them as fair.

7. Concluding Remarks

The study was concluded by examining the results and discussing the main findings in order to respond to the research questions put forth in this research. The three main findings of the research are as follows:

1. Learners' views and feedback on peer assessment in a case study of a MOOC indicated the need for improvement. They specially pointed out the need for instruction and training of peer assessment.
2. It was made clear that the support (i.e. training and instruction) provided during the peer grading activity could have a positive effect on raising the accuracy of peer awarded grades in MOOC.
3. Motivating MOOC learners during the peer assessment activity could encourage them to put conscious effort towards more accurate assessment of their peers' works.

In addition, possible implementations of the findings such as the inclusion of instructional support and motivating MOOC learners in order to navigate them towards conducting quality assessment were discussed. Finally, limitations of the research were pointed out and future research paths such as the need for investigating the effects of similar attempts on increasing the quality of peer assessment across several MOOCs of varying topics were put forth.