A Systematic Analysis of Peer Assessment in the MOOC Era and Future Perspectives

Usman Wahid, Mohamed Amine Chatti, Ulrik Schroeder
Informatik 9 (Learning Technologies),
RWTH Aachen University, Germany
e-mail: {wahid; schroeder}@cil.rwth-aachen.de; chatti@informatik.rwth-aachen.de

***Abstract*— Massive Online Open Courses (MOOCs) have become a cost and time effective choice for learners all across the globe. This has led to new challenges for teachers such as providing valuable and quality assessment and feedback on such a large scale. Recent studies have found peer assessment where learners assess the work of their peers to be a viable and cost effective alternative to teacher/staff evaluation. This study systematically analyzes the current research on peer assessment published in the context of MOOCs and the online tools that are being used in MOOCs for peer assessment. 48 peer reviewed papers and 17 peer assessment tools were selected for the comparison in this study and were assessed on three main dimensions, namely, system design, efficiency and effectiveness. In the light of the comparison and discussion of current research in terms of these categories, we present future visions and research dimensions to improve the peer assessment process in MOOCs.**

Keywords-Open Assessment; Peer Assessment; MOOC; Blended Learning; Peer Reviews, Online Assessment.

I. INTRODUCTION

The advent of Massive Online Open Courses (MOOCs) has revolutionized the field of technology-enhanced learning (TEL). MOOCs enable a massive number of learners from all over the world to attend online courses irrespective of their social and academic backgrounds. MOOCs have been classified in different forms by researchers, including cMOOCs, xMOOCs [1]. cMOOCs allow the learners to build their own learning networks by using blogs, wikis, Twitter, Facebook and other social networking tools outside the confines of the learning platform [2]. Whereas xMOOCs follow a more institutional model, having pre-defined learning objectives e.g., Coursera, edX and Udacity. Apart from these sMOOCs and bMOOCs have also been introduced as variations of the MOOC platform with sMOOCs catering to a relatively smaller number of participants and bMOOCs combining the in-class and online learning activities to form a hybrid learning environment [1].

Irrespective of the classification, MOOCs require their stakeholders to address a number of challenges including and not limited to the role of university/teacher, plagiarism, certification, completion rates, innovating the learning model beyond traditional approaches and last but not the least assessment [3]. Assessment and Feedback are an integral part of the learning process and MOOCs are no different in this regard. However, in the case of MOOCs assessment presents a bottleneck issue due to the massiveness of the course participants and requires increased resources on part of the teachers. This limitation causes many MOOCs to use automated assessments. Peer assessment offers a scalable and cost effective way of providing assessment and feedback to a massive amount of learners where learners can be actively involved in the assessment processes [4]. A significant amount of research is directed towards exploring peer assessment in MOOCs discussing many issues such as the effective integration of peer assessment in MOOC platforms and the improvement of the peer assessment process.

It is evident that peer assessment is a viable assessment method in MOOCs hence, the need for scouting available systems and studies becomes paramount in importance as it could be beneficial for future developments as well as provide a good comparison of available tools. In this study, we look at the state of art in peer assessment in the MOOC era, perceived benefits and challenges of peer assessment. We also look at different tools for peer assessment and the manner in which they try to address the different challenges and drawbacks. The remainder of this paper is structured as follows: Section II introduces peer assessment. Section III is a review of the related work. Section IV describes the research methodology and how we collected the research data. In Section V, we review and discuss the current research based on several dimensions. Section VI summarises the results of our findings. Section VII presents challenges and future perspectives in peer assessment. Finally, Section VIII gives a conclusion of the main findings of this paper.

II. PEER ASSESSMENT

In recent years, student assessments have shifted from the traditional testing of knowledge to a culture of learning assessments [5]. This culture of assessment encourages students to take an active part in the learning and assessment processes [5]. Peer assessment is the flag bearer in this new assessment culture. Peer assessment, is defined by Topping as "an arrangement in which individuals consider the amount, level, value, worth, quality or success of the products or outcomes of learning of peers of similar status" [6].

Peer assessment has been leveraged in a wide range of subject domains over the years [7]. According to Somervell [8], at one end of the spectrum peer assessment may involve feedback of a qualitative nature or, at the other, may involve students in the actual marking process. This exercise may or may not entail previous agreements over criterion. It may involve the use of rating instruments, which may have been designed by others before the exercise, or designed by the user group to meet its particular needs. The use of peer assessment not only reduces the teacher workload; it also

 ISBN: 978-1-61208-471-8

brings many potential benefits to student learning. These benefits include a sense of ownership and autonomy, increased motivation, better learning and high level cognitive and discursive processing [7].

Despite these potential benefits, peer assessment still has not been able to have strong backing from either teachers or students. Both parties have pre-conceived notions of low reliability and validity on their minds when discussing peer assessment [9]. A number of possible factors have been identified for the lack of effectiveness of peer assessment in MOOCs including the scalability issue, diversity of reviewers, perceived lack of expertise, lack of transparency and fixed grading rubrics [10]. The aim of this paper is to examine the available literature and tools for peer assessment, provide a systematic analysis and provide a bigger picture of the research domain.

III. Related Work

Peer assessment in MOOCs is still an emerging field, hence we did not find any research directly related to our work. Luxton-Reily [11] made a systematic comparison of a number of online peer assessment tools in 2009, but the study was conducted with limited dimensions for comparing the tools. The study examined tools including legacy systems, and divided the tools in different categories; namely generic, domain specific and context specific. The study identifies the problem that majority of online tools have been used in computer science courses, and most of the tools could not be used outside the context in which they were developed. This context limitation prevents these tools from being widely adopted which gives rise to the need for more general-purpose tools. Luxton-Reily also stressed the need to investigate the quality of the feedback provided by students [11].

IV. Methodology

The research methodology used for this study is divided in two parts; namely, identification of eligible studies followed by a cognitive mapping approach to find criterion for categorizing and analyzing peer assessment tools.

A. Identification of Eligible Studies

We applied the significant research method of identifying papers from internet resources in our study [12]. This method was carried out in two rounds. Firstly, we conducted a search in 7 major refereed academic databases. These include Education Resources Information Center (ERIC), JSTOR, ALT Open Access Repository, Google Scholar, PsychInfo, ACM publication, IEEEXplorer, and Wiley Online Library. We used the keywords (and their plurals) “Peer Assessment”, “Peer Review”, “Assessment in MOOC”, and “Peer Assessment in MOOC”. As a result, 87 peer-reviewed papers were found. In the second round, we identified a set of selection criteria as follows:

1- Studies must focus on using peer assessment preferably in a MOOC setting.

2- Studies that focus on design of peer assessment systems or that detail the setting in which peer assessment should be carried out were included.

3- Studies focusing on peer assessment in a manual setting were excluded.

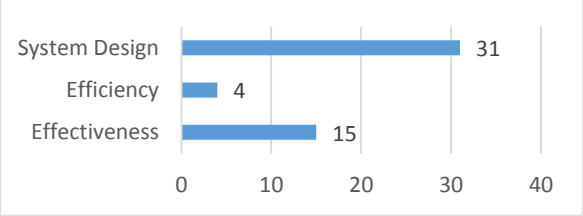
4- Tools older than 10 years have not been included in the study, however, tools having current support are included.

This resulted in a set of 48 research papers/studies on peer assessment in MOOCs and a list of 17 peer assessment tools. These tools include Peer Studio [13], Cloud Teaching Assistant System (CTAS) [14], IT Based Peer Assessment (ITPA) [15], Organic Peer Assessment [16], EduPCR4 [17], GRAASP Extension [18], Web-PA [19], SWoRD (Peerceptiv now) [20], Calibrated Peer Reviews (CPR) [21], Aropä [22], Web-SPA [23], Peer Scholar [24], Study Sync [25], Peer Grader [26] and L²P (Lehr und Lern Portal, RWTH Aachen) Peer Reviews [10]. We also took a look into some open systems that could be used in MOOCs as well, namely: TeamMates [27] and TurnItIn [28].

B. Cognitive Mapping Approach

Cognitive mapping is a method that enables researchers to classify and categorize things into several dimensions based on the research questions [29]. For the sake of our study, we scouted the literature to form a directed cognitive map for each study identifying main ideas that talked about peer assessment. These maps were later analyzed for distinct cluster of concepts, grouping similar terms and ideas. After analyzing the clusters, we were able to identify certain dimensions namely system design, efficiency and effectiveness (see figure 1) which were all part of the discussed peer assessment systems. These dimensions provide an easy and efficient way to assess different peer assessment tools/studies.

FIGURE 1. PEER ASSESSMENT CLASSIFICATION MAP



V. Discussion

This section deals with the critical analysis of the peer assessment literature based on the cognitive mapping dimensions derived in the previous section. In order to capture the information gained from the literature analysis, we partitioned these three categories into ten sub-categories. For the critical discussion part, we discuss the way in which certain tools cater to different dimensions (if at all).

A. System Design

A lot of effort has been put into the design of peer assessment systems, design of certain features provided by the system and the manner in which they are implemented. Nearly 70% of the studies deal in one way or the other with system or a feature design in peer assessment. In the following sections, we discuss some key features of peer assessment systems and the way; different tools realize them.

 ISBN: 978-1-61208-471-8

1) Anonymity: Anonymity is a key feature that is to be kept in mind while designing any peer assessment system, as it safeguards the system against any type of bias (gender, nationlity, friendship etc.) to play a factor in the assessment from peers. There are three levels of anonymity namely, single blind: assessor knows the assessee but the assessee has no idea of the assessor, double blind: both assessor and assessee are unaware of each other and finally no anonymity in which the identity of both the assessor and assessee are known to each other. Most of the systems reviewed in this study follow the principle of double blind reviews to remove bias, however TurnItIn [28] and Study Sync [25] only implement the single blind reviews.Whereas, organic peer assessment [16] has no mention of the feature at all.

2) Delivery: This feature entails the delivery mode of the review, whether it is delivered indirectly (as is the case in most of the MOOC courses), or directly face to face (could be a situation in a bMOOC). All the reviewed systems only support indirect feedback at the moment.

3) Grading Weightage: Almost two third of the reviewed systems assign a pre-defined weightage to the review from the peers in the overall grade. This means that the final grade is calculated by combining the grade from the peers and the instructor and assigning certain weightages to each of them.

Channel: Researchers believe that more reviews help the assessee to have multiple inisghts about their work and learn from them instead of a single point of view being forced upon them. All the reviewed systems provide multi-channel feedback support for the reviews. A study conducted at Stanford and University of California proposed a process of selecting an appropriate number of reviewers needed for each submission by making use of an automated system. Initially, the student grade is predicted by a machine learning algorithm which then estimates the confidence value. This value is used to determine the required number of peer graders [30]. This automated process aims at putting manageable load on peers by trying to reduce the number of peers required for each submission.

Review Loop: The purpose of this feature is to allow the students to work on their assignments in multiple iterations in order to improve the final product and have a better learning outcome. Although, researchers claim it to be a very important feature for any peer assessment tool, only a handful of the reviewed tools actually implement more than one review loops. These systems include PeerStudio [13], EduPCR4 [17], Peerceptiv [20], Aropä [22], Web-SPA [23] and Peer Grader [26]. Peer grader is unique in this respect as it allows for a communication channel between the author and the reviewer to help the authors improve their submissions [26].

4) Collaboration: Collaboration means the ability of the tool to allow students to form and work in small groups. Although many MOOC platforms make use of discussion forums and wikis to enable collaboration and idea sharing between the students, but we found that only three systems actually allow the students to form groups and submit their work in groups. Team mates [27] is an open source tool that allows the students to form smaller groups and submit their work. Also L^2P Peer reviews [10], makes use of a separate module Group Workspace in their learning management system to manage student groups.

B. Efficiency

In this section, we list the features that contribute to the overall efficiency of the system. These features allow the system to be more efficient for its users and help them get the most value out of the system.

1) Feedback Timing: Research has shown that the optimal timing of a feedback is early in the assessment process, as it gives the learners more time to react and improve. Peer Studio, proposes an effective way to reduce the review response time. The learners can submit their work any number of times for a peer review and get the review by reviewing the work of others. A study conducted on the usefulnes of the system concludes that the students in the Fast Feedback condition did better than the No Early Feedback condition group. It also states that on average students scored higher by 4.4% of the assignment's total grade, hence proving the usefulnes of early feedback [13].

C. Effectiveness

Several researchers in TEL have explored how to design effective peer assessment modules with a higher level of user satisfaction. We identified certain features that contribute to the effectiveness of the reviews provided by the peers, which are discussed in the following sections.

1) Rubrics: Rubrics provide a way to define flexible task specific questions that could include descriptions of each assessment item to achieve fair and consistent feedback for all course participants. Studies suggest that asking direct questions for the peers to answer, in order to assess the quality of someone's work enables the reviewer to easily reflect on the quality of submitted work in a goal oriented manner [10]. Hence, a flexible rubric system becomes a must have feature for any good peer assessment system. In our study, we found that majority of the reviewed systems offer this feature in one way or the other with a notable exception of Peer Grader.

A variation of the use of rubrics is the way peer studio tool handles them. The tool allows the teachers to define rubrics and then enforces the students to answer these questions in a better way by using a technique they call scaffolding comments [13]. The system does this scaffolding by making use of short tips for writing comment below the comments box.

2) Validation: A number of studies have been carried out on the validation aspect of the reviews provided by peers, i.e., on methods to make sure that the feedback provided by the peers is valid and of a certain value. Luo et al. [7] conducted a study specifically on Coursera platform to evaluate the

 ISBN: 978-1-61208-471-8

validity of the reviews from peers. In their study they propose that increasing the number of reviewers and giving prior training to the reviewers are some techniques that could be used to bolster the validity of the reviews. Peerceptiv measures the validation of reviews to a submission by simply calculating the agreement rate between different reviewers. It takes score difference, consitency and the spread of scores into consideration for evaluating the validity of reviews [20].

3) Reviewer Calibration: Calibrated peer reviews (CPR) [21] along with some other studies carried out in MOOCs [31] propose a different method to achieve system effectivness, namely, reviewer calibration. In this method, the reviewers are required to grade some sample solutions which have been pre-graded by the instructor to train them in the process of providing reviews.

4) Reverse Reviews: Another interesting method to verify the effectiveness of the reviews is to use the reverse review method. Peer Grader [26] and EduPCR4 [17] tools make use of this method to allow the original authors of the reviewed submissions to rate the reviews they reeived from their peers. The students can specify whether the review helped them in improving their submission, or was of a certain quality, or helped them understand the topic clearly. This review is then taken into consideration at the time of calculation of the final grade, so the peers who provided better reviews have a chance to better their assignment score.

VI. SUMMARY

Table I shows a summary of evaluation of different tools against the dimensions identified in Section IV. The table shows that nearly all the tools reviewed in our study follow a similar system design varying slightly based on the context in which they are used. The only major discrepancy in most tools is their inability to allow students to work in groups (for assignment submission and reviews). Another pattern emerging from studying the table is that more and more tools are giving weightage to the student reviews in the overall grade of the students. This means that the teachers have to be sure about the validity and quality of the student reviews, and the system has to provide features for its insurance. Another useful observation is the usage of assessment rubrics by the tools to help students in the review process. As identified by Yousef et al. [10] rubrics are an easy way to provide learners with task specific questions, allowing achievement of fair and consistent feedback for all course participants.

In the comparison for the validation, we mention all the tools for which a study has been conducted for the validation of peer reviews. It does not specify that the tool have some in-built validation mechanism for the reviews provided by peers. Table I also highlights an important trend in the field of peer assessment for MOOCs. It shows that most systems are moving on from the basic system design and looking for ways to improve the efficiency and effectiveness of the system. This leads to the use of innovative ways to ensure the quality of reviews, and a focus to find ways on improving the overall user experience and learning.

VII. CHALLENGES AND FUTURE VISION

MOOCs with their large number of participants pose a challenge when it comes to assessment and feedback, and peer assessment offers a viable solution to the problem. However, peer assessment itself faces a number of challenges including scalability, reliability, quality and validation. A number of studies have focused on overcoming these limitations, as outlined in the previous section but there is still a lot of room for improvement.

The challenges faced by peer assessment are inherent from the challenges of open assessment in general, and the field of learning analytics offers a number of techniques to overcome these challenges. In this section, we try to offer some solutions from the field of learning analytics, which could be used to overcome some peer assessment challenges.

1) Scalability: The massive number of participants in the MOOC courses requires the feedback provided to students to be scalable as well. This requires the use of certain measures to decrease the time required by the teacher to provide useful feedback to the student submissions. Although peer assessment tries to lessen the teacher's burden but still the teacher has to be in the loop to ensure quality feedback. To overcome this issue of scalability, we could make use of clustering techniques. We could cluster similar submissions

TABLE I. A SYSTEMATIC COMPARISON OF PEER ASSESSMENT TOOLS

	System Design						Efficiency	Effectiveness			
Tools	**Anonymity**	**Delivery**	**Grading Weightage**	**Channel**	**Review Loop**	**Collaboration**	**Time/Rapid Feedback**	**Rubrics**	**Validation**	**Reviewer Calibration**	**Reverse Reviews**
Peer Studio [13]	Double	InDirect	Yes	Multiple	Multiple	No	Yes	Yes	Yes	No	No
CTAS [14]	Double	InDirect	Yes	Multiple	Single	-	No	Yes	Yes	No	No
ITPA [15]	Yes	InDirect	No	Multiple	Single	-	No	Yes	Not measured	No	No
Organic PA [16]	No	InDirect	No	Multiple	Single	-	No	No	Yes	No	No
EduPCR4 [17]	Double	InDirect	Yes	Multiple	Double	-	No	Yes	Not measured	No	Yes
GRAASP extension [18]	No	InDirect	Yes	Multiple	Single	-	No	Yes	Yes	No	No
Web-PA [19]	Yes	InDirect	Yes	Multiple	Single	Yes	No	Yes	Not measured	No	No
SWoRD/Peerceptiv [20]	Double	InDirect	Yes	Multiple	Double	Yes	No	Yes	Yes	No	No
CPR [21]	Double	InDirect	Yes	Multiple	Single	No	No	Yes	Yes	Yes	No
Aropä [22]	Yes	InDirect	Yes	Multiple	Double	-	No	Yes	Yes	No	Yes
Web-SPA [23]	Yes	InDirect	No	Multiple	Double	Yes	No	Yes	Yes	No	No
Peer Scholar [24]	Double	InDirect	Yes	Multiple	Single	No	No	Yes	Yes	No	No
Study Sync [25]	Single	InDirect	No	Multiple	Single	No	No	Yes	Yes	No	No
Peer Grader [26]	Double	InDirect	Yes	Multiple	Double	No	No	No	Yes	No	Yes
L^2P Peer Reviews [10]	Double	InDirect	Yes	Multiple	Single	Yes	No	Yes	Yes	No	No
Team Mates [27]	Double	InDirect	No	Multiple	Single	Yes	No	Yes	Not measured	No	No
TurnItIn [28]	Single	InDirect	No	Multiple	Single	No	No	Yes	Yes	No	No

together and in case of peer assessment, the similar reviews (including rubric answers) could be clustered together to form a single unit. The teacher could easily grade the clusters, in turn saving valuable time.

Another solution to the problem of scalability could be the use of word clouds by extracting important parameters from the submitted work of students. This could help the teacher by providing an overview of the submission and giving a fair idea about the contents. Hence, a teacher could decide if the submission requires in depth review or they could grade based on the provided information.

2) Reviewer Credibility/Reliability: There have been cases identified in peer assessment studies, where students don't take the process of reviewing others work seriously. This leads to invalid reviews and casts a doubt over the credibility of the reviews being provided to students. In this scenario, the teacher must be in the loop to ensure valid reviews. One solution to this could be to rate the reviewers using the reverse review method. This way we could identify possible bad reviewers and screen them out for further reviews or they could be urged to provide better reviews.

3) Validity: We have already seen the usage of calibration to improve the validity of the reviews. Raman and Joachims make use of a statistical method in their study to ensure the validity of the reviews. They use Bayesian ordinal peer grading to form an aggregated ordering for all the submissions in a course room. The difference in ranking from different peers is also taken into account to ensure the effectiveness and validity of reviews [32]. Another approach could be the use of automated assessment, as is the case in automatic essay grading systems. The system takes into account the grade from one human reviewer and the automated assessment grade. If the difference in grades from both sources is greater than a certain threshold, then the system asks for an additional review from a human grader [33]. This technique could be applied to the peer reviews, and if the disagreement between the review from peer and the automated assessment is significant, the system could mark the submission for grading by the teacher or ask for a review from some other peer as well.

4) Quality: Rubrics provide an easy way of improving the quality of the reviews [10]. The peer assessment system could enhance this by providing a way for the teacher to specify common mistakes that students make, so that the reviewer could look for these and in turn improve review quality.

5) System Configuration: Another improvement to the peer assessment tools could be to allow the user to configure different settings from a central location rather than making it a part of system design. Majority of peer assessment systems in use today have pre-defined configuration in features like anonymity, review loops, grading weightage etc. These pre-configured settings make it difficult for the tool to be used in a more generic way and in different contexts. Hence, a tool that allows its users to configure all these settings could be a lot more useful across different domains and have a higher acceptance rate from users across the glo.

VIII. CONCLUSION

Peer assessment is a rich and powerful assessment method used in technology-enhanced learning (TEL) to improve learning outcomes as well as learner satisfaction. In this paper, we analysed the research on peer assessment published in the MOOC era, and the tools that could be used to provide peer assessment capabilities in a MOOC. A cognitive mapping approach was used to map the selected studies on peer assessment into three main dimensions namely, system design, efficiency and effectiveness. The following is a summary of the main findings in our study as well as aspects of peer assessment that need further research, according to each dimension.

A. System Design

The analysis of the peer assessment research showed that majority of the systems are designed on similar lines to each other, differing in only a small number of features or the way these features are implemented. Despite these possible differences in implementation, the general idea for different system features remains the same across different tools. However, several features concerning system design need a better acceptance across these tools: (1) Collaboration: The tools should allow the students to work in a collaborative environment and submit their assignments and even review in groups. (2) Review Loops: In our opinion, all peer assessment tools should provide at least double review loops to give students more chances of improvement and in doing so we leverage the peer assessment model in a better way to achieve positive results.

B. Efficiency

Studies have established the positive effect of timely feedback on student performance but the assessment tools are lagging far behind in this regard. In our opinion, tools should focus on efficient ways to decrease feedback time, and focus on more innovations to make the process more efficient.

C. Effectiveness

Several methods are being used in peer assessment to increase effectiveness of the reviews and in turn the learners' satisfaction with peer assessment. Although, rubrics, reviewer calibration and reverse reviews are good ideas to improve the effectiveness of the reviews; more research has to be put into measuring the validity of the reviews provided by peers. Future research needs to find new ways to record validity of reviews and improvements to this validity. The systematic comparison also reveals certain patterns and trends across the analysed tools. It points out the fact that most tools are quite similar in system design, and the way they carry out the peer assessment process. The difference arises in the way the apply validation and effectiveness techniques to the peer reviews. The study also highlights the

 ISBN: 978-1-61208-471-8

shift in focus from basic system design to innovative ways of improving the quality and effectiveness of the reviews provided by peers.

The study concludes with providing a list of open challenges in the peer assessment process/systems and proposes certain techniques that could be applied to address these challenges. The proposed solutions include a number of techniques from the field of learning analytics including statistics, visualizations, and data mining techniques that could prove useful in improving the peer assessment process/tools.

REFERENCES

[1] G. Siemens, "MOOCs are really a Platform. Elearnspace (2012)." 2011.

[2] G. Siemens, "Connectivism: A learning theory for the digital age," 2014.

[3] A. M. F. Yousef, M. A. Chatti, U. Schroeder, M. Wosnitza, and H. Jakobs, "MOOCs-A Review of the State-of-the-Art," in *Proc. CSEDU 2014 conference*, 2014, vol. 3, pp. 9–20.

[4] R. O'Toole, "Pedagogical strategies and technologies for peer assessment in Massively Open Online Courses (MOOCs)," 2013.

[5] A. Planas Lladó, L. F. Soley, R. M. Fraguell Sansbelló, G. A. Pujolras, J. P. Planella, N. Roura-Pascual, J. J. Suñol Martínez, and L. M. Moreno, "Student perceptions of peer assessment: an interdisciplinary study," *Assess. Eval. High. Educ.*, vol. 39, no. 5. 2014, pp. 592–610.

[6] K. Topping, "Peer assessment between students in colleges and universities," *Rev. Educ. Res.*, vol. 68, no. 3, 1998, pp. 249–276.

[7] H. Luo, A. C. Robinson, and J.-Y. Park, "Peer grading in a mooc: Reliability, validity, and perceived effects," *Online Learn. Off. J. Online Learn. Consort.*, vol. 18, no. 2, 2014.

[8] H. Somervell, "Issues in assessment, enterprise and higher education: The case for self-peer and collaborative assessment," *Assess. Eval. High. Educ.*, vol. 18, no. 3, 1993, pp. 221–233.

[9] O. McGarr and A. M. Clifford, "'Just enough to make you take it seriously': exploring students' attitudes towards peer assessment," *High. Educ.*, vol. 65, no. 6, 2013, pp. 677–693.

[10] A. M. F. Yousef, U. Wahid, M. A. Chatti, U. Schroeder, and M. Wosnitza, "The Effect of Peer Assessment Rubrics on Learners' Satisfaction and Performance within a Blended MOOC Environment," in *Proc. CSEDU 2015 conference*, 2015, vol. 2, pp. 148–159.

[11] A. Luxton-Reilly, "A systematic review of tools that support peer assessment," *Comput. Sci. Educ.*, vol. 19, no. 4, 2009, pp. 209–232.

[12] A. Fink, *Conducting research literature reviews: from the Internet to paper*. Sage Publications, 2013.

[13] C. Kulkarni, M. S. Bernstein, and S. Klemmer, "PeerStudio: Rapid Peer Feedback Emphasizes Revision and Improves Performance," in *Proceedings from The Second (2015) ACM Conference on Learning@ Scale*, 2015, pp. 75–84.

[14] T. Vogelsang and L. Ruppertz, "On the validity of peer grading and a cloud teaching assistant system," in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, 2015, pp. 41–50.

[15] K. Lehmann and J.-M. Leimeister, "Assessment to Assess High Cognitive Levels of Educational Objectives in Large-scale Learning Services," 2015.

[16] S. Komarov and K. Z. Gajos, "Organic Peer Assessment," in *Proceedings of the CHI 2014 Learning Innovation at Scale workshop*, 2014.

[17] Y. Wang, Y. Liang, L. Liu, and Y. Liu, "A Motivation Model of Peer Assessment in Programming Language Learning," *arXiv Prepr. arXiv1401.6113*, 2014.

[18] A. Vozniuk, A. Holzer, and D. Gillet, "Peer assessment based on ratings in a social media course," in *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, 2014, pp. 133–137.

[19] P. Willmot and K. Pond, "Multi-disciplinary Peer-mark Moderation of Group Work," *Int. J. High. Educ.*, vol. 1, no. 1, 2012, p. p2.

[20] J. H. Kaufman and C. D. Schunn, "Students' perceptions about peer assessment for writing: their origin and impact on revision work," *Instr. Sci.*, vol. 39, no. 3, 2011, pp. 387–406.

[21] P. A. Carlson and F. C. Berry, "Calibrated peer review/sup TM/and assessing learning outcomes," in *fie*, 2003, pp. F3E1–6.

[22] J. Hamer, C. Kell, and F. Spence, "Peer assessment using arop{ä}," in *Proceedings of the ninth Australasian conference on Computing education-Volume 66*, 2007, pp. 43–54.

[23] Y.-T. Sung, K.-E. Chang, S.-K. Chiou, and H.-T. Hou, "The design and application of a web-based self-and peer-assessment system," *Comput. Educ.*, vol. 45, no. 2, 2005, pp. 187–202.

[24] S. Joordens, S. Desa, and D. Paré, "The pedagogical anatomy of peer-assessment: Dissecting a peerScholar assignment," *J. Syst. Cybern. Informatics*, vol. 7, no. 5, 2009.

[25] D. L. White, "Gatekeepers to Millennial Careers: Adoption of Technology in Education by Teachers," *Handb. Mob. Teach. Learn.*, p. 351, 2015.

[26] E. F. Gehringer, "Electronic peer review and peer grading in computer-science courses," *ACM SIGCSE Bull.*, vol. 33, no. 1, 2001, pp. 139–143.

[27] G. Goh, X. Lai, and D. C. Rajapakse, "Teammates: A cloud-based peer evaluation tool for student team projects," 2011.

[28] S. Draaijer and P. van Boxel, "Summative peer assessment using 'Turnitin' and a large cohort of students: A case study," 2006.

[29] S. McDonald, K. Daniels, and C. Harris, "Cognitive mapping in organizational research In C. Casssell & G. Symon (Eds.), Essential guide to qualitative methods in organizational research" London: Sage, 2004, pp. 73-85.

[30] C. E. Kulkarni, R. Socher, M. S. Bernstein, and S. R. Klemmer, "Scaling short-answer grading by combining peer assessment with algorithmic scoring," in *Proceedings of the first ACM conference on Learning@ scale conference*, 2014, pp. 99–108.

[31] J. Wilkowski, D. M. Russell, and A. Deutsch, "Self-evaluation in advanced power searching and mapping with google moocs," in *Proceedings of the first ACM conference on Learning@ scale conference*, 2014, pp. 109–116.

[32] K. Raman and T. Joachims, "Bayesian Ordinal Peer Grading," in *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, 2015, pp. 149–156.

[33] H. Chen and B. He, "Automated Essay Scoring by Maximizing Human-Machine Agreement.," in *EMNLP*, 2013, pp. 1741–1752.

 ISBN: 978-1-61208-471-8