Organic Peer Assessment

Steven Komarov
Harvard SEAS
Cambridge, MA USA
komarov@seas.harvard.edu

Krzysztof Z. Gajos
Harvard SEAS
Cambridge, MA USA
kgajos@eecs.harvard.edu

CHI 2014, April 26 – May 1, 2014, Toronto, Ontario, Canada.
ACM 978-1-4503-2473-1/14/04.

Abstract
We present a web-based tool for a type of peer assessment we dubbed *organic.* In organic peer assessment there are no upper or lower limits on the number of assignments each peer has to review, avoiding the common issue of prematurely coercing students into activities they might fear and dislike. Instead, peer assessment occurs as a side effect of activities students find intrinsically motivating. We outline the basic set of functionality required for the implementation of our vision for peer assessment in an online environment and present the results of a preliminary study we conducted in a flipped classroom. We found that the quality of the summative assessment produced by the peers matched that of experts, and we encountered strong evidence that our peer assessment implementation had positive effects on achievement. We conclude with a discussion arguing that organic peer assessment is a valuable technique—distinct from formal peer assessment—for deployment in MOOCs.

Author Keywords
Collaborative Learning; Peer Assessment; Flipped Classroom; Online Education;

ACM Classification Keywords
H.5.3 [Group and Organization Interfaces]: Computer-supported cooperative work

Introduction

In recent years peer assessment has grown in popularity in both formal and online classroom settings, especially in the areas of higher education [1] [2]. The growth can be explained by the teaching community's increased understanding of its benefits, which fall into two broad categories: pedagogical and logistical.

From the pedagogical perspective, peer assessment is a valuable instructional tool that helps students develop skills in critical enquiry, reflection, and "learning to learn" [3]. It has been shown to facilitate deeper learning and to imrove students' metacognitive understanding [4]. Students who engage in peer assessment also develop valuable workplace skills such as the ability to collaborate and the ability to evaluate the work of other professionals [1].

From the logistical perspective, peer assessment makes grading of assignments scale. Students can receive more feedback on their work as the number and length of assignments is no longer exclusively constrained by the availablility of the teaching staff. This is especially relevant to MOOCs and other courses where the student body vastly outnumbers the available graders. While automated grading is a common alternative for such courses, it may be hard to implement and severely limits the types of assignments that can be used. Peer assessment offers more flexibility and has already been implemented on a massive scale with encouraging results [2].

Student Aversion to Peer Assessment

However, despite the advantages it brings to the table, peer assessment suffers a major drawback. Many educators report that students often show aversion towards the acts of assessing or being assessed by their peers [4] [5]. The reasons for this resistance are complex and include lack of reliability, lack of validity, low perceived expertise and confidence, cultural aspects of the student-teacher power relation, inexperience, and others [4]. We believe that despite the potential pedagogical benefits of peer assessment, coercing students head-first into an activity many of them strongly dislike and fear might lead to a long-lasting negative outlook if done prematurely.

Organic Peer Assessment

To address the issue of student aversion, we experimented with a form of peer assessment we named *organic*. There are three key characteristics of organic peer assessment: (1) no strict limits on participation, no deadlines; students can choose to review as many or as few assignments as they feel comfortable, including none. (2) All forms of peer feedback, including grades, play no role in determining the students' official course grades. (3) The peer assessment happens in the background, as a byproduct of students performing activities they find intrinsically valuable.

We tested our vision for peer assessment in a flipped classroom at a traditional institution. The resulting rate of participation and the quality of the assessments produced by the peers exceeded our expectations. Furthermore, students who actively engaged in peer assessment showed significant academic gains as reflected in their exam performance. While our study was conducted in a flipped, but nevertheless formal classroom setting, the tool and the workflow around it can readily be used in an online course.

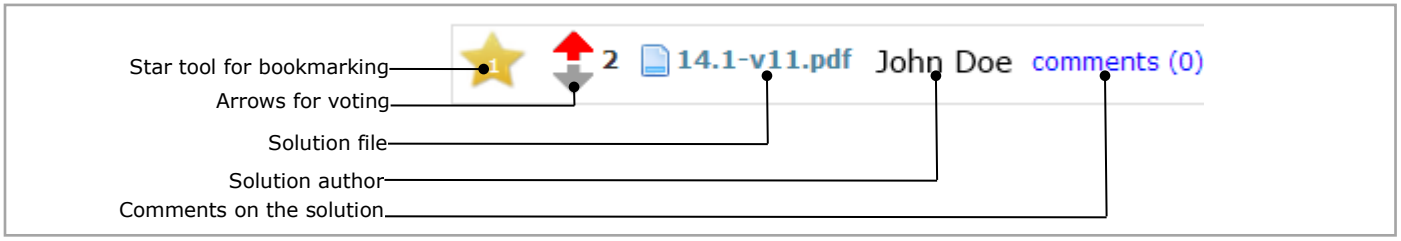


Figure 2: Tools for organic peer assessment in the bookmarklet

Figure 1: Solutions tab displaying a list of 5 solutions. Student names are blurred.

Organic Peer Assessment Tool

We dubbed the ad-hoc peer assessment tool we created for the course the "collaborative bookmarklet", because it was implemented as a browser bookmarklet. When activated the tool allows students to vote, bookmark, and comment on solutions submitted to the course website in the form of PDF files. The tool achieved this by modifying, on startup, the underlying code of the course website displaying the list of class-wide viewable student-submitted solutions. Thus students who chose to use the bookmarklet gained access to additional peer assessment and learning functionality, while those who chose not to, saw a plain list of PDF files. The bookmarklet's functionality is organized in three tabs.

Solutions Tab

The solutions tab displays a list of all solutions submitted by students for a given module (Figure 1). There is a high redundancy of submitted solutions – a single problem typically receives at least 3 distinct solutions. Each solution has a row of tools for peer assessment (Figure 2).

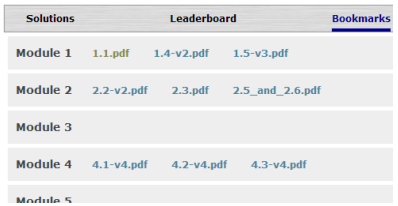
Leaderboard Tab

The tab displays a ranking of all students in the class according to two scores. The peer score reflects the feedback received by a student, while the community service score reflects the student's frequency of participation. Scores below the 40th percentile are not publically visible to avoid potential embarrassment.

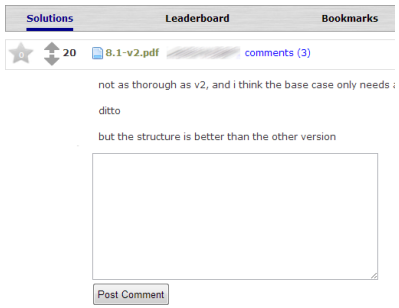
Bookmarks Tab

The bookmarks tab displays all the solutions bookmarked (starred) by a student one a single page with solutions from the same module (lesson) grouped together. In other words, if a student bookmarks the solution on the solutions page, it will appear on the bookmarks page. The purpose of the tool is to allow

Leaderboard tab



Bookmarks tab



Solution with a comment thread containing three comments.

students to quickly access solutions they have selected for later study.

Students can perform three basic peer-assessment actions with the bookmarklet tool: bookmarking (starring), voting, and commenting.

VOTING

Voting is the primary mechanism for peer feedback. It is equivalent to assigning a grade. Each student can upvote a solution up to 3 times or downvote a solution up to 3 times. Thus a student can assign a grade between -3 and 3 points. Students upvote and downvote using the arrows to the left side of the solution file (Figure 1). The number displayed next to each arrow indicates the total points assigned by all peers and it is visible by all class members.

BOOKMARKING (STARRING)

The star tool allows the student to mark a solution that he wants to review in the future. The number inside of the star indicates how many times this solution has been bookmarked by other students and serves as an indicator of quality – the higher the count, the better the solution. Each solution starred by a student appears on his or her bookmarks page.

When a student votes on a solution he does it for the benefit of his peers; when he stars a solution he does it for his or her own benefit. Thus, even though both bookmarking and voting serve as a quality indicator, the motivation behind them is different.

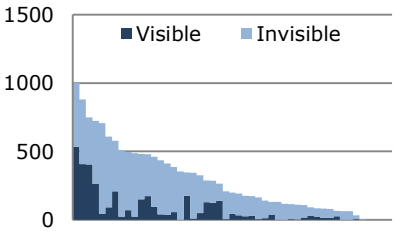
COMMENTING

Students can post comments to a solution to ask for clarification or simply provide feedback. The comments are anonymous to encourage participation, as anonymity makes the atmosphere more informal. However on some occasions students can choose to sign their name. This happens frequently when the author of the solution is responding to feedback. There is a separate comment thread for each solution.

Study

We conducted an informal study of organic peer assessment in the course CS20 "Discrete Mathematics for Computer Science" offered at Harvard University in the spring of 2013. This was an active learning course following the "flipped classroom" methodology. At home, students were preparing ahead of each class meeting by reading a textbook and watching recorded lectures. In class, they were spending the majority of time working in small groups of five, solving problems assigned by the instructors. The problems were at a medium-to-high difficulty level for an introductory course, and students typically solved no more than two during the 60 minutes allotted for group work.

Because the course staff did not have the resources to formally review the solutions of all small-group problems in addition to grading regular assignments and exams, some students did not receive corrective feedback on their work and underperformed on the first midterm exam. To address this we created a simple assessment tool that allowed students to review and learn from each other's work. While students were required to post their solutions to a class-wide viewable webpage weekly, using our tool to collaboratively evaluate those solutions was purely optional. However, students had a strong incentive use to review each other's solutions, because the exams in the course were based on the same problems. Also, we hoped that



Number of actions per student. The top 20% of students performed over 80% of all assessment (visible) actions.

the leaderboards will create a fun, game-like atmosphere encouraging participation.

Results

Participation

The majority of students enrolled in the course (45 out of 49) used the bookmarklet at least once during the seven weeks of the course it was available. The students submitted 503 solutions for peer review and carried out 1611 visible assessment actions (203 comments, 421 bookmarks, and 987 votes). Additionally there were over 12 thousand invisible actions such as reading comments, viewing points, accessing leaderboards and bookmarks.

Validity

To measure the level of agreement between expert-assigned scores and student-assigned scores, four Course Assistants (CAs) independently graded a randomly chosen subset of 18 solutions among all solutions that received at least two student votes.

Two agreements were computed: (1) the inter-rater agreement between CAs using the Krippendorf's alpha statistic, and (2) the inter-rater agreement between CAs and students, also using the Krippendorf's alpha statistic. For the latter only the average score of each respective group was used. The agreement among the four CAs was $\alpha=0.73$ and the agreement between CAs and students was $\alpha=0.72$. The Pearson correlation between CA and student grades (votes) was $r=0.82$, $N=18$, $p<0.01$. A correlation of $r=0.82$ is much stronger than correlations observed in experiments with enforced peer assessment [6].

Effects on achievement

We uncovered strong link between the use of the bookmarklet and achievement gains. For our analysis we divided the students into two equally-sized groups: active students (used the bookmarklet most frequently) and inactive students (used the bookmarklet least frequently). The active students showed the largest improvement in exam score after the introduction of the bookmarklet in the course (Figure 3). The difference between the two groups was statistically significant: $t(43)=2.2$, $p<0.04$.

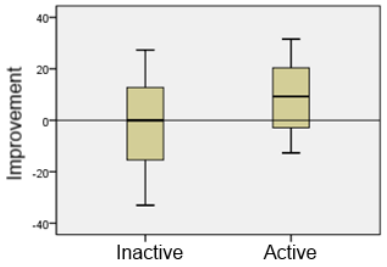


Figure 3: Improvement in exam score for active and inactive users of the bookmarklet.

Organic Peer Assessment at Scale

The success of our first experiment with organic peer assessment encouraged us to test our vision at a MOOC scale in the future. We believe that organic peer assessment holds a number of benefits for MOOCs.

First, it could improve the reliability and validity of peer feedback, because it does not force incapable students to produce assessments. This is especially relevant to MOOCs with unrestricted enrollment that may have a large number of students with insufficient skills and preparation.

Second, organic peer assessment intends to encourage unsupervised exchange of ideas and informal discussions similar to the discussions students sitting next to each other in a classroom can have. This could bring MOOC students one step closer to having a small classroom experience.

Third, based on the results of our study, it can be argued that organic peer assessment can lead to learning gains. Furthermore, it likely produces learning that is different, possibly in a good way, from the learning in enforced peer assessment. Thus, it can potentially enhance the pedagogical value of MOOCs.

Fourth, we believe that the no-deadline, no-limits approach of organic peer assessment is the best way to introduce yet another activity to an already overworked group of learners. This appears even more relevant in MOOCs with a substantial number of employed adults and other overcommitted students who might find it too burdensome to keep up with the additional deadlines of enforced peer assessment.

Finally, based on the results of our study, we believe that organic peer assessment has the ability to motivate enough participation and produce a significant volume of assessments despite its voluntary nature. In MOOCs such assessments can offset some of the need for human graders and complement existing automated grading systems.

Acknowledgments

This work was supported by the National Science Foundation under fellowship number DGE-1144152 and by the Alfred P. Sloan Research Fellowship. We would like to thank Paul Bamberg, Chris Dede, and the staff of CS20 for their assistance.

References

[1] K. J. Topping, "Peer Assessment," *Theory Into Practice,* pp. 48:1, 20-27, 2009.

[2] C. Kulkarni, K. P. Wei, H. Le, S. R. Klemmer and e. al., "Peer and self assessment in massive online courses," *ACM Transactions on Computer-Human Interaction (TOCHI),* vol. 20, no. 6, p. Article No. 33, 203.

[3] D. C. R. S. J. Boud, "Peer Learning and Assessment," *Assessment and Evaluation in Higher Education,* pp. 413-427, 99.

[4] O. McGarr and A. M. Clifford, "Just enough to make you take it seriously: exploring students' attitudes towards peer assessment".

[5] N.-F. Liu and D. Carless, "Peer feedback: the learning element of peer assessment.," *Teaching in Higher Education,* vol. 11, no. 3, pp. 279-290, 2006.

[6] L. J. A. Bouzidi, "Can Online Peer Assessment be Trusted?," *Educational Technology and Society,* pp. 257-268, 2009.