Effective Student Crowdpolling: Key Decisions and a Learning Case

Full Paper

Earl McKinney
Bowling Green State University
Emckinn@bgsu.edu

Bethany Niese
Bowling Green State University
Bniese@bgsu.edu

Abstract

While student peer evaluations have been used in education to generate grades and feedback for students, little work has been done on using students in a crowdpolling context to evaluate other student work. In this paper we highlight student crowdpolling in general and provide a learning case from our own experience. We also develop a list of key decisions and challenges for using student crowdpolling.

Keywords

Crowdsourcing, crowdpolling, student peer evaluation

Introduction

Crowdsourcing of grading (also called peer evaluation) by students is becoming more popular with new learning management system (LMS) options and growing class sizes. Class sizes have been steadily increasing and will continue to do so in the foreseeable future (Messinger 2010; Pomponio 2016; Ryman 2012). One concern with increasing classes is the increase in demand for the instructor to provide feedback and evaluation for all the students (Ryman 2012). As class size increases, it becomes more challenging to assign assignments with rich feedback as grading is too laborious. A potential solution to this problem is to leverage student peer evaluation—to engage students in grading and providing feedback to their fellow students.

Crowdsourcing

Crowdsourcing is a methodology of merging feedback from many people (Raman & Joachims 2014). It has many applications such as analyzing free-form web-based content, understanding online marketplace behavior, and various online verification tasks (Ipeirotis & Paritosh 2011). Examples of using crowdsourcing in education is student peer evaluation, having students lead discussion board interactions, asking students for questions for a guest speaker, and having students generate a course long project such as a list of security audit items in an IS security class.
This paper draws a distinction between crowdsourcing and crowdpolling. Crowdsourcing involves providing verbal or written feedback. Crowdpolling is a simplified version of crowdsourcing where the only feedback provided is a score such as A, B, or C or “above expectations”, “meets expectations”, or “below expectations”.

Where student crowdpolling has been used, empirical evidence in general shows little statistical differences between student-provided evaluations and instructor-provided evaluations (de Alfaro & Shavlovsky 2013; Kritikos, Woulfe, Sukkar, & Saini 2011; Luaces, Diez, Alonso-Betanzos, Troncoso, & Bahamonde 2017; Raman & Joachims 2014; Wright, Thornton, & Leyton-Brown 2015). This result has been consistent across numerous studies. The comparison between student-provided evaluations and instructor-provided evaluations has occurred in real-time (compared evaluations submitted by these two groups in the same course at the same time) (de Alfaro & Shavlovsky 2013; Kritikos et al. 2011; Luaces et al. 2017; Raman & Joachims 2014) and when comparing grades for the same course during years when crowdsourcing was used versus when it was not (Wright et al. 2015).

While some progress has been made with crowdsourcing/peer evaluations, there has been little empirical evaluation of student crowdpolling in education.

### Student Crowdpolling in Education

A student crowdpolling process requires students to vote on a learning activity accomplished by another student or student group. Students could vote on best presentation, vote on the top three posts on a discussion board, provide an ordinal ranking of group projects, choose the three best questions a guest speaker should address, rank student summarizes of chapters, or pick the best student generated example of a course topic.

There are many different ways to approach crowdsourcing which causes variation in the way studies are conducted. Technology can be used to collect the scores and calculate final grades (Albano, Capuano, & Pierri 2017; de Alfaro & Shavlovsky 2013; Luaces et al. 2017; Robinson, 2001; Wright et al., 2015) or manual methods can be used (Raman & Joachims 2014). The task can involve assigning categorical groups (Albano et al. 2017; de Alfaro, Shavlovsky, & Polychronopoulos 2016; Sadler & Good 2006), or ordinal ranking (Raman & Joachims 2014).

While these student crowdpolling studies and assignments vary, based on our own experience and these preliminary studies, we believe effective student crowdpolling is based on a number of key decisions made by the instructor. These decisions are listed in Figure 1 and then described.

<table>
<thead>
<tr>
<th>Assignment type</th>
<th>Level of Instructor/TA involvement</th>
<th>Student grader guidance</th>
<th>Grading type</th>
<th>Grade calculation</th>
<th>Student motivation</th>
</tr>
</thead>
</table>

**Figure 1. Key Decisions**

### Key Decisions

One of the key decisions is assignment type, which assignment should this technique be applied to. In most cases, LMS tools are used to grade simple question and answer methodologies such as true and false questions and multiple choice (Noorbehbahani & Kardan 2011). Effective student crowdpolling depends on learning to use this technique on appropriate assignments. Students can be asked to grade a variety of assignment types which include presentations (Kritikos et al. 2011), and written assignments such as essays, short answers, and papers (de Alfaro & Shavlovsky 2013; Luaces et al. 2017; Robinson 2001; Wright et al. 2015).
Effective Student Crowdpolling

The second key decision is the level of instructor and TA involvement in the student peer evaluation method chosen. There are student peer evaluation methods which involve a high degree of participation from the instructor and GAs. This participation can be in terms of providing student grader training (Wright et al. 2015), involvement in settling a student complaint involving the grade they received (Kritikos et al. 2011), providing rubrics and gold standards (de Alfara et al. 2014), and ensuring high-quality feedback and grades have been provided (deAlfara et al. 2014).

The third key decision is the amount of guidance and training the students receive. Numerous studies have proposed solutions to student peer evaluations which involve a very detailed rubric to act as guidance to student graders and/or grader training (de Alfaro & Shavlovsky 2013; Kritikos et al. 2011; Luaces et al. 2017; Raman & Joachims 2014; Wright et al. 2015). Some provide badges or levels to the students who pass the training and/or are assessed to have provided high-quality grades (de Alfaro & Shavlovsky 2013). Studies have shown that students provide higher quality evaluations when they have been trained on grading as opposed to simply being given a rubric (Albano et al. 2017; Robinson 2001; Sadler & Good 2006). However, training usually involves either a sophisticated technology or time and effort from the instructor and TAs, which may not be feasible depending on the class size and the instructor’s capacity.

The fourth key decision is the type of grading the students are asked to provide. Crowdpolling can also be used which is the act of providing a cardinal (categorical) evaluation such as A, B, C, D, or F (Luaces et al. 2017; Raman & Joachims 2014). Another way that students can be asked to provide feedback is to perform an ordinal ranking of other students’ work (Luaces et al. 2017; Raman & Joachims 2014). Finally, students can be asked to simply select the best or the top n from a group of assignments their peers turned in.

Studies have found that ordinal grading is easier to provide and require less information input to students than cardinal feedback (Ipeirotis & Paritosh 2011; Raman & Joachims 2014). Studies have not agreed on how these two methods compare in terms of quality. They have been found to have similar effectiveness (Raman & Joachims 2014) and that ordinal methods are more reliable (Luaces et al. 2017). It has also been found that effectiveness depends on the amount of quality graders that are in a course. Specifically, when “lazy graders” are added to the model, ordinal methods outperform the cardinal methods (Raman & Joachims 2014). Students have been shown to prefer cardinal grading to ordinal ranking, so there is less resistance when cardinal grading is assigned (Albano et al. 2017; de Alfaro & Shavlovsky 2013; de Alfaro et al. 2016).

The method used to calculate grades is also an important decision. Grades have been calculated in many ways. Most of the grade calculations for student peer evaluations involve a combination of the grade(s) assigned to the student and the amount and/or effectiveness of their grading of other students (de Alfaro & Shavlovsky 2013; Luaces et al. 2017; Robinson, 2001). The level of precision and accuracy can be determined by instructors and TAs or by technical tools. Wright et al. (2015) did not include the aspect of grading in the final grade, but that study involved a small class which may need less incentive to perform the evaluation.

Finally, it is important to decide how students will be motivated. Students may initially be reluctant as they may feel inadequate or that the task is too unusual. Instructors need to consider an incentive opportunity for students. Some examples of incentives are to reward students whose grades are similar to other student evaluations; however this technique often leads to limited use of the entire grading scale.

Benefits and Challenges

Student crowdpolling can have many benefits. The primary benefit is that instructors can assign meaningful assignments and not have the entire responsibility for completing all of the grading. This is especially of value in large lecture classes which makes it infeasible for instructors and teaching assistants (TAs) to assign and grade assignments. Students can also benefit from this methodology. When students are more engaged, attentive, reflective, and analytical when they are accountable for reviewing other’s work (Kritikos et al. 2011; Luaces et al. 2017) quality feedback and student learning can be improved. It also provides them with opportunities to build skills such as critical thinking and giving and receiving feedback with peers (Robinson 2001).
There are many challenges to implementing and maintaining a student crowdpolling project. Many of them are related to the student’s self-efficacy. Many students feel unqualified to assess others’ work fairly, both in terms of how to provide feedback and/or grades and in terms of mastery of the content (Kritikos et al. 2011). This leads to discomfort in completing the grading exercise as well as the extent to which they feel that the grade they received is accurate and fair (Kritikos et al. 2011; Raman & Joachims 2014). Students have also been observed to lack objectivity. They tend to award higher grades to friends and resist providing low grades or negative feedback to any of their classmates (Kritikos et al. 2011).

The actual or perceived ability of the students to provide high-quality feedback also remains a question (Raman & Joachims 2014). Many student peer evaluation methods include an aspect of student training for effective grading and/or grading monitoring. However, this can involve highly complex technology that not all instructors have access to or the skills to build (Ipeirotis & Paritosh 2011; Noorbehhahani & Kardan 2011; Pérez-Marín, Pascual-Nieto & Rodríguez 2009) or it may involve a massive effort by the instructor and TAs. If the students aren’t trained in grading or monitored, the result may be a lower-quality evaluation or the perception of a lower-quality evaluation (Raman & Joachims 2014). The result of this could involve student concern and subsequent complaints that may undermine the effectiveness of the class.

Another challenge is student participation. Many student peer evaluation methods depend on more than one peer grade so participation in the grading activity is critical. Even when providing grades and/or feedback to their peers is a part of their final grade, not all students consistently complete the grading which makes final grade calculation difficult or impossible.

The final challenge is related to the amount of instructor and TA involvement in the student peer evaluation method. Many instructors implement peer evaluation methods to save time. However, implementing and maintaining the student peer evaluation takes time and resources (Kritikos et al. 2011; Wright et al. 2015). When planning and designing the student peer methodology, a close eye needs to be on the amount of student grader training, monitoring, and grader quality evaluation occurs. If there are too many students to allow for this choice, then the amount of time needed by the instructor and TAs may become overwhelming and they become bottlenecks in the grading process.

Learning Case

At our university we have developed and used a student crowdpolling exercise we call the Current Event game. During each week of the semester, student teams submit web articles to a custom made web site that makes the article submitter anonymous and records grades of that article made by classmates. When classmates grade articles the web site is shown in Figure 2:

![Image of the Current Event game interface]

**Figure 2: 6 point grade and comment**
On the form shown in Figure 2 the grading student marks one of the 6 available grades, and answers the question posed in the description line by the submitter in the comment box. Each article is graded by every class member (except class members who are on the submitter’s team). An overall score is created at the end of grading week by calculating an average grade on a 6 point scale.

**Syllabus**

To better understand the context for the game, here is the syllabus description of the game:

One key aspect of IT is rapid change. To stay up to date with new technologies it is important to read about current events and developments in the field. To do this we will play a current events game. Students submit websites of interesting current events about IT to the course website. These current events are read by all students prior to class, each student giving each site a grade. For each site submitted an overall grade is calculated from all the individual grades given by the students.

Each student is assigned to a submitting team, and each team is a member of one of two alliances. Students do not know which alliance their team is assigned to. Most weeks four teams will submit. After the submission deadline, all the students in the class read each of the sites and grade each one before class. When current events are discussed in class, the submitting team explains its submission, and leads a discussion of the responses.

When grading sites, students should try to use the entire grading scale. Significant deviations from this distribution of grades will lower your current event grade.

**Objectives**

There are four objectives of this game. First, improve understand of current technology—students learn from the web articles interesting current information technologies that are having an impact on business. Second, develop presentation skills—students practice leading a discussion in class on an article. Third, gain first-hand experience delivering a learning product—students experience the challenge of providing a service (an interesting web article) to a customer group—their fellow students. Fourth, learn technology—teachers can use the software and database technologies supporting the game to help explain IT topics in class.

**Evaluation**

Students are evaluated in four ways. The most important is how well a team scores when it submits a site. A second input is how many sites a student grades before class meets to discuss that site. A third input is the variety of grades used by a student. The final input is based on how frequently a student responds to the questions asked on the site.

**Feedback/Results**

We have used this game for a number of years in classes ranging from freshmen to executive MBAs, in accounting and IT courses, and with several instructors. We have changed the way the game is played recently, so the most recent quantitative results are limited. We have not yet conducted a specific quantitative assessment of this current version of the game.

However, we can report preliminary results based on an end of course survey we use in all courses. At the end of each term we ask students to evaluate the major course activities. Activities include computer lab sessions, guest speakers, textbook, tutorials, presentations etc. The prompt is: “This activity helped me understand course concepts.” In the five most recent courses the game has been played it scored in the top half of activities for each of the five sections. The only other activity listed in the top half on each of the five sections was exams. While a much more specific assessment is planned, at least the initial results have been very positive.
Effective Student Crowdpolling

Benefits

While the following benefits are provided for the Current Event game, we believe most apply to all recurring student crowdpolling activities.

One benefit of the game is for students to have first-hand experience with a crowdpolling event. Often their first reaction is to question the validity of the grades, but after several reassuring weeks, the conversation can turn to more substantive topics on crowdpolling such as how to construct the rules, importance of trust in the system and the concept, and how to control bad behavior such as overly negative grades or no variation in grades.

A second benefit is that students learn how to listen to consumers of a product or service. A very common question from the students is to ask the teacher for the criteria. They will ask, “What type of article should we submit.” We emphasize we’d like to see articles only on technology (not politics or sports) that your fellow students would appreciate reading. We admit that when we play the game and submit articles, we often do not do well. The message is that it is hard to anticipate your audiences’ tastes, welcome to business. Also, we emphasize that each week we show the scores of all articles so that future submitters can learn from the “sale” of earlier articles and find ones that will score well.

A third benefit is improving students’ ability to find information that other young people will appreciate. This is not the common academic assignment of producing your own information to be scored by a professor who is not a peer or in your generation. Students do not produce the information they find an article, and the consumer is a young demographic.

A fourth benefit is that the software and database generated by playing the game are excellent examples of several topics in an Introduction to IS, Database, Systems Analysis and Design, Analytics or Web Development course. Even if a custom made crowd polling exercise like the Current Event game is not used, any common crowdpolling (Kickstarter, Prediction Markets, etc.) example that is used can provide the same type of examples.

We use the game in the Introduction to IS class to explain 3 tier architecture, controls on web forms, and databases. In the Database course, we have students make an E/R diagram of the data used to support the game, and to write SQL statements that would be helpful for the instructor and students. In a Design class we use the game as a prototype and have the class, after playing the game for a month, redesign it using the systems development steps. In Analytics we use previous semesters data on visualization software like Tableau to look for patterns in the data, to contrast OLTP and OLAP, and to make a denormalized data structure. Finally, in a Web Development course the game could be used to build a mobile front end or some other platform.

The final two benefits are for the instructor. By using a crowdpolling exercise the instructor can learn about the interests of their audience, and gather useful articles and examples to use in upcoming course topics. It can also reduce workload as the teacher is not doing all the evaluation.

Finally, we have found it helpful to use as a role play device in class. After playing the game, the teacher tells the student that he/she is a consultant hired by the school to improve the game/activity. Students become end users and collaboration between end users and systems designers can be demonstrated.

Challenges

Students object to having their grade determined by other students. To reassure students we remind them the game is never worth more than 10% of their grade. We also emphasize that final scores will have minimal variation which also reduces the impact of the game on their final grade.

Students often fail to grade articles on time or submit on time. We have learned to be consistent throughout the term. Every week, half the teams submit, the next week the other half. At one time in the past we mixed up the schedule so that different teams would be competing against different opponents every week. We also used to not require submissions during exam weeks or other breaks in the schedule. We now mechanically do half each week to minimize the demands on students’ memory.
Conclusion

Student crowdpolling has potential to generate useful feedback for students and reduce grading by instructors, but work needs done. First, evaluations of this method need to be more specific about improvements in learning outcomes. Researchers need to also evaluate other factors that may impact success such as prior student experience with student crowdpolling, type of course, educational level, and specific educational outcomes.

REFERENCES

de Alfaro, L., & Shavlovsky, M. 2013. CrowdGrader: Crowdsourcing the evaluation of homework assignments.
Pomponio, R. 2016. Universities face rising admissions yields and growing class sizes. The Daily Pennsylvanian
Ryman, A. 2012. Bigger Arizona college class sizes a growing trend. AZCentral