Tournament Approach to Peer Review in a Quantitative Course

Nicole Michelotti, Jared Tritz, David Winn, and Tim McKay
Department of Physics, University of Michigan

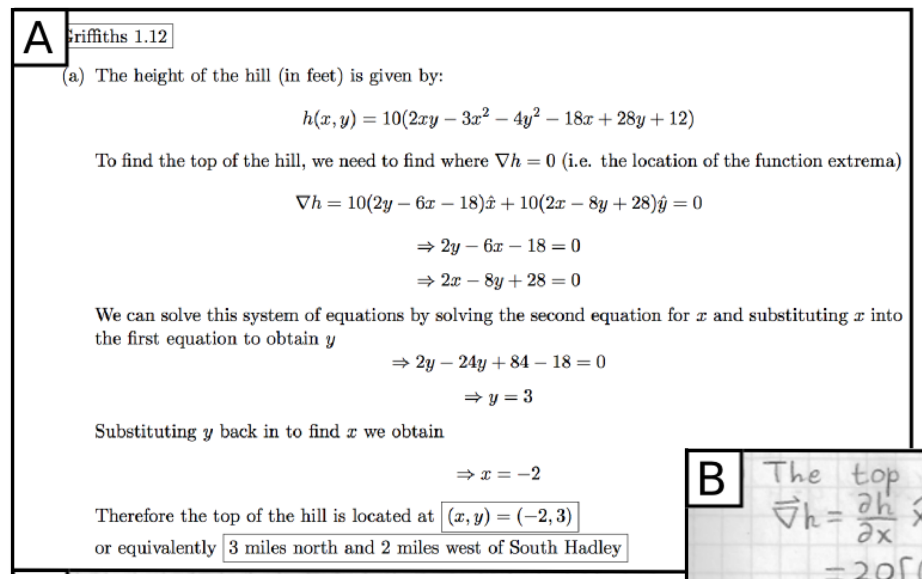
Introduction

The educational practice of peer evaluation is widely used in writing intensive courses, where students exchange papers in class and use criteria described by a rubric to evaluate the writing of their fellow classmates. There are studies that support that students learn through assessing their peers [1,2]. Despite the wide use of peer evaluation in writing intensive courses, quantitative courses have not adopted this mode of learning. While most assigned problems have one solution, for complicated problems posed in upper-level courses, some ways of approaching a problem are more sophisticated than others. It is conceivable that the same benefits of being exposed to an evaluating the work of their peers that students receive from writing assignments would translate to benefits obtained from evaluating their peers' problem-solving assignments. Here, we describe how we implemented peer evaluation in an upper-level physics course and results.

PROTon

For peer review in this course, we used a homebuilt Peer Review Online Tournament (PROTon) system. Students uploaded their work into the online system (A,B). Then students were presented with pairs of assignments three times, each time being asked which assignment in the pair was better than the other. They were also given the option of providing written feedback. Then, the students were shown their rank in the class based on the feedback of their peers (C).

A


Griffiths 1.12

(a) The height of the hill (in feet) is given by:

$$h(x,y) = 10(2xy - 3x^2 - 4y^2 - 18x + 28y + 12)$$

To find the top of the hill, we need to find where $\nabla h = 0$ (i.e. the location of the function extrema)

$$\nabla h = 10(2y - 6x - 18)\hat{x} + 10(2x - 8y + 28)\hat{y} = 0$$

$$\Rightarrow 2y - 6x - 18 = 0$$

$$\Rightarrow 2x - 8y + 28 = 0$$

We can solve this system of equations by solving the second equation for x and substituting x into the first equation to obtain y

$$\Rightarrow 2y - 24y + 84 - 18 = 0$$

$$\Rightarrow y = 3$$

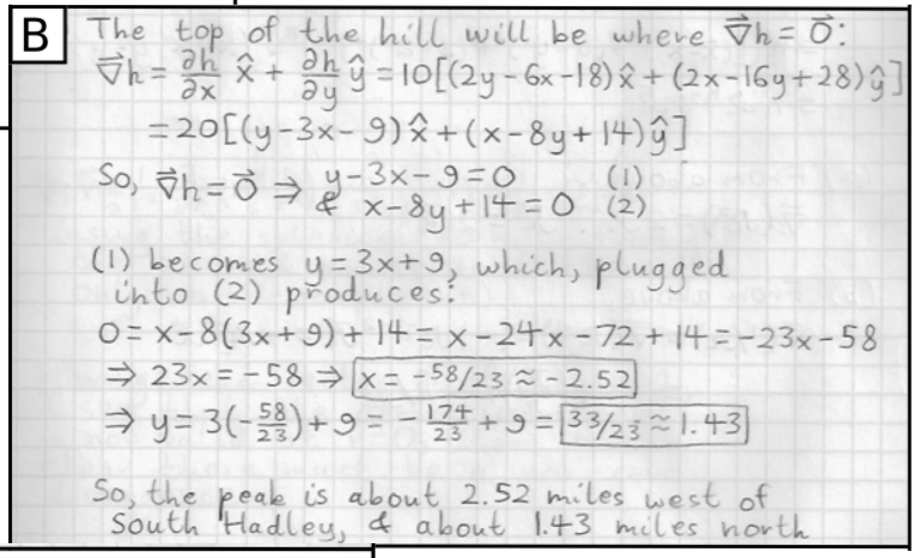
Substituting y back in to find x we obtain

$$\Rightarrow x = -2$$

Therefore the top of the hill is located at $\boxed{(x,y) = (-2,3)}$
or equivalently $\boxed{\text{3 miles north and 2 miles west of South Hadley}}$

A) Example of a student submission in the top half of the rankings.

B


The top of the hill will be where $\vec{\nabla} h = \vec{0}$:

$$\vec{\nabla} h = \frac{\partial h}{\partial x}\hat{x} + \frac{\partial h}{\partial y}\hat{y} = 10[(2y-6x-18)\hat{x} + (2x-16y+28)\hat{y}]$$

$$= 20[(y-3x-9)\hat{x} + (x-8y+14)\hat{y}]$$

So, $\vec{\nabla} h = \vec{0} \Rightarrow y - 3x - 9 = 0$ (1) & $x - 8y + 14 = 0$ (2)

(1) becomes $y = 3x+9$, which, plugged into (2) produces:

$$0 = x - 8(3x+9) + 14 = x - 24x - 72 + 14 = -23x - 58$$

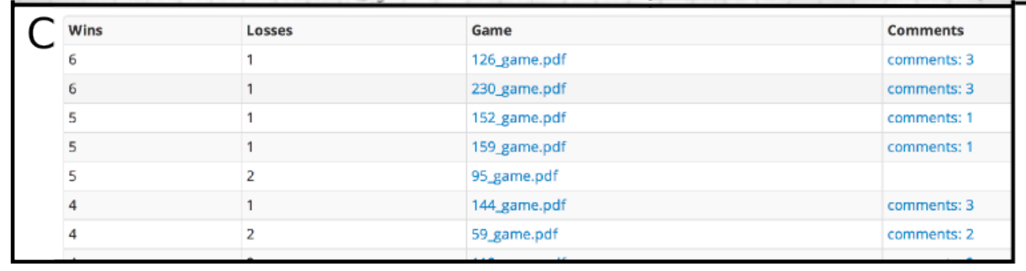
$$\Rightarrow 23x = -58 \Rightarrow \boxed{x = -58/23 \approx -2.52}$$

$$\Rightarrow y = 3\left(-\frac{58}{23}\right) + 9 = -\frac{174}{23} + 9 = \boxed{33/23 \approx 1.43}$$

So, the peak is about 2.52 miles west of South Hadley, & about 1.43 miles north

B) Example of a student submission in the bottom half of the rankings.

C

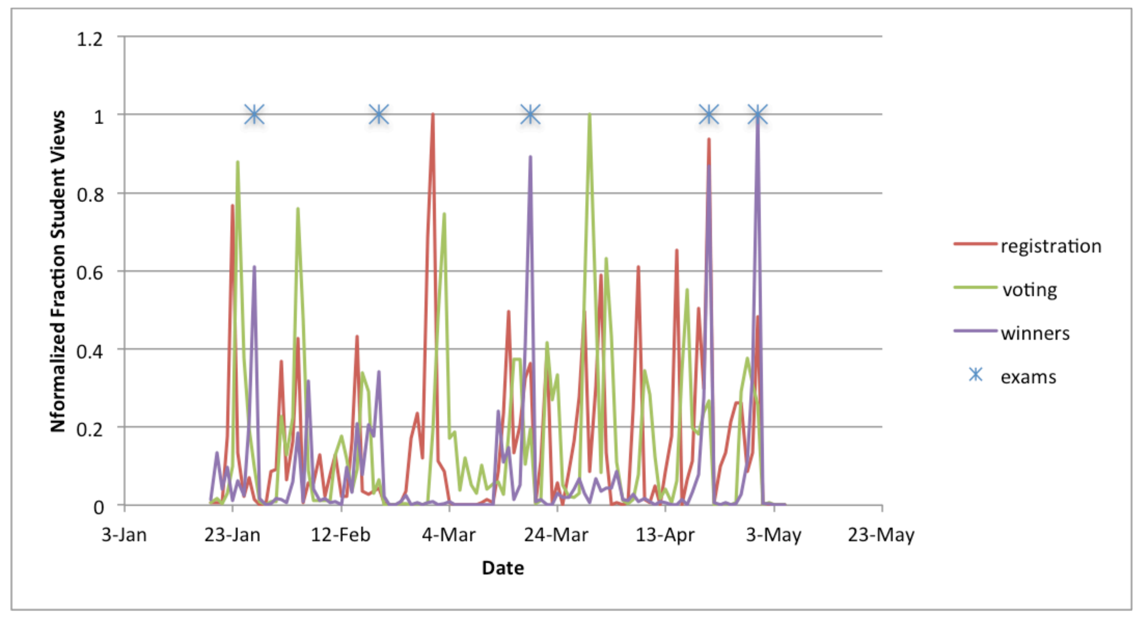


Wins	Losses	Game	Comments
6	1	126_game.pdf	comments: 3
6	1	230_game.pdf	comments: 3
5	1	152_game.pdf	comments: 1
5	1	159_game.pdf	comments: 1
5	2	95_game.pdf	
4	1	144_game.pdf	comments: 3
4	2	59_game.pdf	comments: 2

C) Student view of seeing the result of the tournament once it is finished.

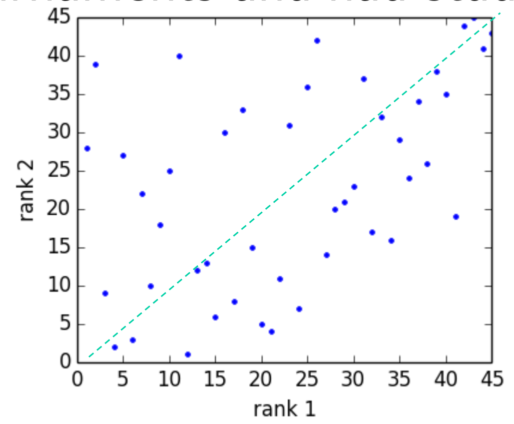
Usage

Students used the system in the way we expected. Namely, the submitted their assignment, voted, and then observed others. As the semester progressed, fewer students observed the other students until immediately before an exam. Then, they used the assignments to study for the exam.



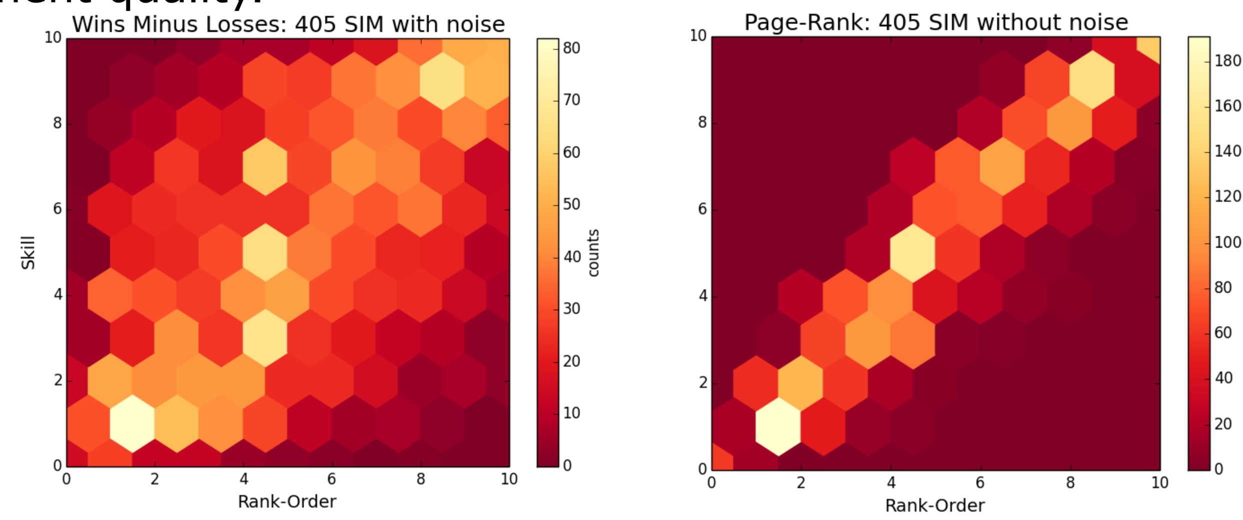
Accuracy in Rankings

One way we looked at accuracy in ranking was we cloned four of the tournaments and had students vote twice. Below is an example of one.



Rank (top or bottom)	Top 5, Rank 1	Bottom 5, Rank 1	Top 5, Rank 2	Bottom 5, Rank 2
1	Incorrect	Correct	Correct	Incorrect
2	Incorrect	Incorrect	Correct	Incorrect
3	Correct	Incorrect	Correct	Incorrect
4	Correct	Incorrect	Correct	Incorrect
5	Incorrect	Incorrect	Correct	Incorrect

Rank 1 is the first time the students go through the tournament and rank 2 is the second time they go through the tournament with the same questions. If the two were the same, all the data would lie along the green line which it doesn't. Looking at the content of the solutions, we see that the accurate ones didn't always make it to the top in the first round, but did in the second. While anecdotal, this suggests that students may have learned the correct solutions through the process of voting. As further confirmation that students not knowing the solutions may impact the error, students were provided with solutions the second half of the semester. The average number of ranking violations dropped from 14% to 11%, a 20% improvement with a 4 sigma significance. Further analysis of the voting quality suggests that only 66% of the votes aligned with true document quality (left below) and the expected improvement (right below) if 100% of the votes align with true document quality.



Student Feedback

At the end of the semester, we asked the students to fill out a survey to give us feedback on the PROTon and their experience of peer evaluation in a physics course. 23 of 48 students in the course completed the survey.

Overall, students liked having peer evaluation in class with 74% of them recommending that we continue using the system in class. The primary benefits they listed were learning from others, learning from writing explanations, and creating community among students.

> *"Knowing that my classmates are grading it, I try to make my solutions as easy to follow as possible. This means I have to think more about how I'm wording explanations and how I layout the solution. Often this extra thought helps me understand the topic better."*

The negative feedback we received included unfair evaluations, the process was time consuming, and demotivating.

> *"Allows me to get away with slacking off and not putting a ton of effort and still get points for submission and voting. And even without the initial effort, I still learn the material and the problems by seeing others students' submissions."*

The suggestions we received for future iterations included incorporating grading into the system and assigning fewer problems for peer evaluation.

> *"I think not all problems are worth doing peer-review tournament. Some homework problems are quite basic and straightforward. They are essential part of learning and understanding the materials, but not really worth time looking at and voting for almost-everyone-gets-it-correctly solutions. But on the other hand, the more challenging ones are really worth doing the peer-reviews."*

Conclusions

Peer evaluation can be used in quantitative courses effectively provided that the problems are challenging enough that there are multiple approaches to achieving the same correct solution. To achieve more accurate rankings, solutions in some form and a rubric should be provided.

Future Directions

The grading scheme was primarily based on whether the student participated and voted. It would be worth rethinking the grading scheme to encourage ideal student learning behavior. It would also be beneficial to have only the problems that would have the greatest learning gains included in the tournament and to require more of the students when ranking such as grading and requiring written feedback.

Acknowledgments & References

Special thanks to the students of Physics 405 who participated in this study.

1. K. Topping, Review of Educational Research **68**, 249 (1998).
2. K. Lundstrom and W. Baker, Journal of Second Language Writing **18**, 30 (2009).