

Subject: TP Msg. #1538 Peer Evaluation of Class Participation
From: Rick Reis <reis@stanford.edu>
Date: 1/30/17, 11:41 AM
To: "tomorrows-professor@lists.stanford.edu" <tomorrows-professor@lists.stanford.edu>



This memo describes a mechanism for evaluating class participation in courses where it matters, refined and developed over a couple of decades but surely not perfected.

1538 Peer Evaluation of Class Participation

Folks:



The posting below describes an interesting approach to evaluating class participation in courses. It is by Michael O'Hare and it was posted on April 16, 2016 on the Berkeley Center for Teaching & Learning website (<http://teaching.berkeley.edu>). This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License 2013 (may be reproduced with credit for non-commercial purposes. Michael

O'Hare's blog: <http://www.samefacts.com/>

Regards,

Rick Reis

reis@stanford.edu

UP NEXT: The Slow Graduate Student: Do Less and Be Mindful



[Subscribe/Unsubscribe](#)

[TomProf Online](#)

**Connect with
us on Social
Media**

Tomorrow's Teaching and Learning

-----1,626 words -----

Peer Evaluation of Class Participation

This memo describes a mechanism for *evaluating class participation* in courses where it matters, refined and developed over a couple of decades but surely not perfected.

Background

Plenary sessions of a college or graduate course are increasingly regarded as an opportunity for the students to apply and explore the tools and content of the curriculum collaboratively (cf “Flipped Classroom” literature, eg <http://www.knewton.com/flipped-classroom/>

Examples include “case-method” teaching (business or law school versions), multiple small-group discussions with coaching, and any discussion-based classroom pedagogy. This memo does not review the many reasons nor contexts in which learning is best accomplished as a collaborative, social process.

Why?

In this learning environment, students teach each other extensively, and a course grade (to be fair, and to send the right signals) must in part reflect their respective success at doing this. Furthermore, college and graduate students have been socialized extensively to believe that flattering the prof will be good for them, and commonly fear that course grades are zero-sum, so they can advance only at the expense of others. A lot of learned behavior and expectations have to be undermined.

I also take as given that the correct criterion for this part of a grade is *student’s contribution to the learning of others*. The problem this criterion presents is that it cannot be observed outside the heads of those doing the learning, and proxy indicators like “how often student A’s contributions match what I (the prof) believe to be correct” are compromised by my ego and anyway say nothing about the value those contributions are, or are not, creating for other students.

On the principle that I have the right to demand information (like answers on exams) that I need to make a fair performance evaluation, I demand information about their learning from others. In response to student preference that I grade class performance—that they are diffident about ‘grading each other’—I’m happy to say “of course I assign this grade, like any other grade. But you have to give me the information with which I can do so.” I suggest to students who believe they can learn without others that they will do better on the web and in the library, and not to take the course.

Finally, as Lauren Resnick pithily observed, “in school, collaboration is cheating; in the workplace, it’s essential”. When I write letters of recommendation, I often have occasion to include the following text, and I think it has a good effect:

Student X took my course Y in semester Z [paper, projects, yada yada] In this course, class participation counts for 00% of the grade [varies from 25% to 40%] and is assessed by the other students in a confidential survey. Wallflowers, unprepared students, and air hogs tend to do poorly on this element. X received a CP grade [of G/in the top 00% of the class], and I consider this an indicator of real leadership potential.

Survey

Twice during the semester and a third time at the end, I circulate an Excel spreadsheet with two alphabets of named rows, distinguished by color (example in Appendix A). The students are instructed to record a

score from 1 to 5 for each other student, in the second alphabet for (i) students in their section (ii) students who critiqued their paper drafts (iii) students from whom they received a draft to critique, and in the first for everyone else. Scores must total 3N for a course of N students. They give themselves 7 in the second alphabet.

A few wiseguys occasionally give everyone the same grade: I discard their scores as uninformative (which costs them their own 7). I calculate the mean scores for each student in each panel (I usually have a GSI copy the score column from each response into a master spreadsheet), weight the means in the second alphabet 1.5 to 2x the scores in the first, and sum them into a total score. I order the names by total score, *alphabetize the names within quartiles* (or terciles, depending on the size of the class—no reason someone should be at the very top or very bottom of a list like this), and publish the resulting list (without numerical scores).

The first two rounds of this survey don't count for grades, but are purely advisory; the last one counts. With the GSI's (teaching assistants), I assign a letter grade to the student receiving the lowest score, and the other grades go up from there to A or A+. In principle, and often nearly in fact, everyone can get a very high grade.

Ancillary practices

Some standard discussion management practices should be recalled here, because a grading system is not independent of other elements of a course's culture.

A lecture hall is not a discussion classroom. Students need to see each other's faces, and need name cards, every day, with names on the front *and back*. Early in the semester, it's necessary to bring markers and card stock to class daily, and invite every student without a name card down to the front of the class to make herself another, in a nice way. You need to learn the names. A good trick for this is to go around the room with a video camera one day, having each student hold up his name card and say his name and 'one interesting fact about yourself'; a few times through this tape and you will have the names cold.

Peer CP grading requires additional practices to signal and reinforce what's sought and why, and repetition is very important here. For example, post the video described above on the course website for the students to use. The idea that students are responsible for the learning of others is not a model they slip into easily. I like to emphasize the devious incentives to help others improve their performance built into the grading rule.

Everything about this peer grading process must be transparent to the students, except who gave whom what score. All the information in this memo is shared with the students, sometimes more than once, in the syllabus and in class, from the start of the course.

I do not require attendance. I tell the students that they are grownups, and if they have more to gain being somewhere else they should definitely be there. But I record attendance carefully and include it on the survey form, with instructions to use it as they wish, or not at all. Attendance in courses graded in this fashion hovers around 95%; this is probably higher than my own lifetime "showing up for important stuff" rate. I also tell them that everyone misses a class now and then, but if they have to it's polite to email the class explaining why so people don't get the wrong idea. They are quite reluctant to do this. Some still email me for permission to miss class, and just I tell them they're not there for my benefit and I'm not in charge of that "permission" either way.

I tell the students that grownups use laptops in important meetings and that others will probably be happy if they find something on the web in the middle of class that advances the discussion (this is fairly common). They figure out on their own pretty quickly that reading email or bidding on Ebay with their

name card right in front of them, and three or four students beside and behind them who will be grading their CP, is not a good idea, and I have not had any problem with laptops or phones in class.

Students sometimes ask me for criteria to use in scoring. I tell them that unlike an exam, there's no reason everyone's performance should be assessed the same way by everyone, as people learn in different ways and use different gifts to teach. But I distribute a list of criteria students have found useful, especially to broaden their scope (Appendix B).

Almost immediately, someone grouching about the process will use language like "I don't want to/feel able to evaluate other students." I lie in wait for this, and pounce on it, in a nice way: "Of course you shouldn't do that: no human being has the right to evaluate another person! This is about evaluating *performance at a specific task*. But one reason performance evaluation is the most corrupt and incompetent function in almost any organization, and why people hate it and avoid it, is that it *feels like* evaluating people. So we have to be careful not to use that language, even as shorthand."

I distribute the first two survey results with some language reminding people that every group of people including Cy Young Award winning pitchers, Nobel Prize physicists, and even Berkeley students, has a top, middle, and bottom tercile on any given measure. I offer "what can I do with this information?" advice, in the email, along the lines of "first, look at the people in the top tercile. What are they doing? Try it! Second, look at the people lower in the list and see what you can do to encourage them to get in the game more. Third, pick three people at random from the class, and take each one out for a latte in return for telling you two things they think you could do to get a higher score, and two things you should do even more of." I also cold-call people in the bottom tercile or bottom two quartiles, because plain shyness is often one component of not contributing, and a fair amount of this still water actually runs pretty deep. I never cold-call enough and students are always grateful when I do it more.

Appendix A- CP Form

Appendix B- Peer Evaluation Criteria

"Desktop faculty development 100 times per year."

Over 62,000 subscribers at over 950 institutions in more than 100 countries

TOMORROW'S PROFESSOR eMAIL NEWSLETTER

Archives of all past postings can be found here:

Sponsored by the Office of the Vice Provost of Teaching and Learning at Stanford University

NOTE: To SUBSCRIBE or UNSUBSCRIBE to the Tomorrows-Professor List click HERE

--+***==--+***==--+***==--+***==--+***==--+***==--+***==

tomorrows-professor mailing list

tomorrows-professor@lists.stanford.edu

<https://mailman.stanford.edu/mailman/listinfo/tomorrows-professor>

