

## When Justice Peeks

### Evaluating Students in Case Study Teaching

*Clyde Freeman Herreid*

In mid-December, in the wake of the great chad debate when the fate of the U.S. presidency had not yet been decided, politicians and demonstrators were in the streets and in front of the camera calling for “justice!” The courtroom had replaced the ballot box in deciding who would win and who would lose. While the patient public waited for judicial wisdom and a sensible conclusion, we all hoped for a nonpartisan ruling—the sort of judgment one would expect from and impartial judiciary. This is the ideal, isn’t it?

Somewhere around every courthouse in America, I suppose you can find the symbol of that legal ideal: a statue of a robed woman holding the scales of justice in one hand and a sword in the other to smite the transgressors of the law. The lady of justice comes to us moderns from the pre-Hellenic pantheon. She is Themis, advisor to Zeus, who became an oracle at Delphi and was known as a goddess of divine justice. There are other pretenders to the scales and the sword, including the Egyptian goddess Maat and the Roman goddess Justitia, but I stand with the Greeks on this one. Themis it is, and she is without corrup-

tion, avarice, prejudice, or favor. She is blindfolded. This is to ensure that she is impartial and fair. But can she be fair without peeking?



The thing I most dislike about teaching is assigning grades. In one breath I know what I mean when I do it—John gets an A and Sarah a C. In the next breath I think it ridiculous, for I do not know what John and Sarah know or what they grasped. Moreover, both John and Sarah may actually think I know what I am doing, yet hate me for it. And maybe I deserve it—sometimes.

Evaluation is an odious affair. I can even remember a time when faculty refused to do it wholesale. During the Vietnam War when pubescent youths were conscripted into the armed forces if they didn’t maintain their grades, there were faculty who abandoned grading altogether, giving all-comers A’s so they would not have “blood on their hands.” They couldn’t abide sending young men into battle because of their academic misdeeds. Instead, other men died in their place. We are still laboring with the grade inflation that resulted.

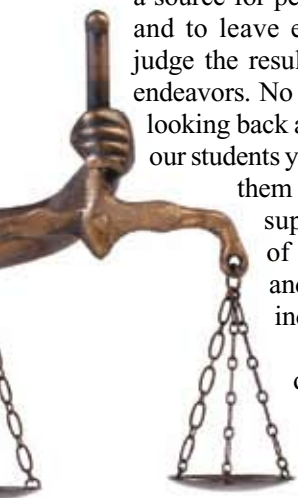
I recall the arguments of those turbulent days. We teachers should forsake evaluation altogether. Our job should be solely to teach. We argued with a certainty that only young faculty can muster. Our role was to act as a library—

---

*Clyde Freeman Herreid, editor of JCST’s “Case Study” department, is Distinguished Teaching Professor in the department of biological sciences, State University of New York, Buffalo, NY 14260-1300; e-mail: herreid@acsu.buffalo.edu.*

---





a source for people to gain wisdom and to leave enriched. Let others judge the result of our educational endeavors. No more this Janus face looking back and front, welcoming our students yet looking away from them when they need our support. The dual role of faculty as instructor and evaluator is odious indeed.

The voices of those days are gone, and we have not shed the onerous duty of evaluation. Given our unhappy lot, what are we to do with this unenviable task, especially when we are thrown into the unfamiliar and discomfiting role of evaluating students in a case study class? If we think there is uncertainty in the normal classroom, what an imbroglio lies in wait for us in the case study classroom.

How do you grade students in classes with case teaching? There are a host of possibilities. I'm only going to deal with a couple. Let's start with the toughest.

♦ **Evaluating class discussion.** Business school case teachers do it all the time. It's not uncommon for them to base the final course grade on 50 percent class participation. And this with 50-70 students in a class! This sends shudders up the spines of most science teachers. Yet, what's so tough about the concept? We are constantly making judgments about the verbal statements of our colleagues, politicians, and even administrators. Why can't we do it for classroom contributions?

Most of our discomfort comes from the subjective nature of the act, something that we scientists work hard to

avoid in our work-a-day world. It may be that we are even predisposed to become scientists because we are looking for a structured and quantifiable world. Flowing from this subjective quandary is the fact that we feel we must be able to justify our grades to the students. We are decidedly uncomfortable if we can't show them the numbers. This is one of the reasons that multiple-choice questions have such appeal for some faculty.

But let's take a look at how the business school people evaluate case discussion. Some of them try to do it in the classroom, making written notes even as the discussion unfolds, using a seating chart, and calling on perhaps 25 students in a period. As you might expect, this usually interferes with running an effective discussion. Other instructors tape-record the discussion and listen to it later in thoughtful contemplation. Most folks, however, sit down shortly after their classes with seating chart in hand and reflect on the discussion. They rank student contributions into categories of excellent, good, or bad, or they may use numbers to evaluate the students from 1 to 4 with 4 being excellent. They may give negative evaluations to people who weren't prepared or were absent. These numbers are tallied up at the end of the semester to calculate the grade. And that's as quantified as it gets.

I especially like mathematician/philosopher Blaise Pascal's view of evaluation: "We first distinguish grapes from among fruits, then muscat grapes, then those from Condrieu, then from Desargues, then the particular graft. Is that all? Has a vine ever produced two bunches alike, and has any bunch produced two grapes alike?"

"I have never judged anything in exactly the same way," Pascal continues. "I cannot judge a work while doing it. I must do as painters do and stand back, but not too far. How far then? Guess . . ."

♦ **Ask for a product.** The simplest solution to case work evaluation is to forget classroom participation and grade everything on the basis of familiar criteria, say exams and papers. This puts professors back in customary territory without the sweat of novelty. Even business and law school professors use this strategy as part of their grades. I'm all for this. In fact, I always ask for some written analysis in the form of journals, papers, and reports. Along with an exam, these are my sole bases for grades. I don't lose sleep over evaluating class participation.

A word about exams. You can give any sort of exam in a case-based course, including multiple choice, but doesn't it make sense to have at least part of the exam a case? If you have used cases all semester and trained students in case analysis, surely you should consider a case-based test. Too often we test on different things than we have taught.

♦ **Peer evaluation.** Some of the best case studies involve small group work and group projects. In fact, I strongly believe teaching cases this way is the most user-friendly for science faculty and the most rewarding for the students. Nonetheless, even some aficionados of group work don't like group projects. They say, how do you know who's doing the work? Even if they ask for a group project, they argue against grading it. They rely strictly on individual marks for a final grade determina-

***Just as in courts of law where there are conflicting personalities, evidence, and judgment calls, teachers carry their own prejudices and experiences into the classroom.***

tion. I'm on the other side of the fence. I believe that great projects can come from teams, and if you don't grade the work, what is the incentive for participating? Moreover, employers report that most people are fired because they can't get along with other people. Not all of us are naturally team players. Practice helps. So, I'm all for group work including teamwork during quizzes where groups almost invariably perform better than the best individuals. But we have to build in safeguards like peer evaluation.

"Social loafers" and "compulsive workhorses" exist in every class. When you form groups such as those in Problem-based Learning (PBL) and Team Learning (the best ways to teach cases, in my judgment), you must set up a system to monitor the situation. In PBL it is common to have tutors who can make evaluations. Still, I believe it is essential to use peer evaluations. I use a method that I picked up from Larry Michaelsen in the School of Management at the University of Oklahoma.

At the beginning of every course I

explain the use of these anonymous peer evaluations. I show students the form that they will fill out at the end of the semester. Then they will be asked to name their teammates and give each one the number of points that reflects their contributions to group projects throughout the course. Say the group has five team members—then each person would have 40 points to give to the other four members of his team. If a student feels that everyone has contributed equally to the group projects, then he should give each teammate 10 points. Obviously, if everyone in the team feels the same way about everyone else, they all will get an average score of 10 points. Persons with an average of 10 points will receive 100 percent of the group score for any group project.

But, suppose that things aren't going well. Maybe John has not pulled his weight in the group projects and ends up with an average score of 8, and Sarah has done more than her share and receives a 12. What then? Well, John gets only 80 percent of any group grade and Sarah receives 120 percent.

There are some additional rules that I use. One is that a student cannot give anyone more than 15 points. This is to stop a student from saving his friend John by giving him 40 points. Another is that any student receiving an average of seven or less will fail my course. This is designed to stop a student from doing nothing in the group because he is simply trying to slip by with a barely passing grade and is willing to undermine the group effort.

Here are some observations after many years of using peer evaluations:

- ♦ Most students are reasonable. Although they are inclined to be generous, most give scores between 8 and 12.
- ♦ Occasionally, I receive a set of scores where one isn't consistent with the others. For example, a student may get a 10, 10, 11, and

**Table 1. Peer Evaluation Form**

Name _____	Group # _____
<p>This is an opportunity to evaluate the contributions of your teammates to group projects during the semester. Please write the names of your teammates in the spaces below and give them the scores that you believe they earned. If you are in a group of five people, you each will have 40 points to distribute. You don't give yourself points. (If you are in a group of four, you'll have 30 to give away. In a group of six you'll have 50 points, etc.) If you believe that everyone contributed equally to the group work, then you should give everyone 10 points. If everyone in the group feels the same way, you all will receive an average of 10 points. Be fair in your assessments, but if someone in your group didn't contribute adequately, give them fewer points. If someone worked harder than the rest, give them more than 10 points.</p> <p>There are some rules that you must observe in assigning points:</p> <ul style="list-style-type: none"> <li>♦ You can not give anyone in your group more than 15 points.</li> <li>♦ You do not have to assign all of your points.</li> <li>♦ Anyone receiving an <i>average</i> of less than seven points will fail the course.</li> <li>♦ Don't give students a grade that they don't deserve.</li> </ul>	
Group members	Score
1. _____	_____
2. _____	_____
3. _____	_____
4. _____	_____
5. _____	_____
6. _____	_____
Please indicate why you gave someone less than 10 points.	
Please indicate why you gave someone more than 10 points.	
If you were to assign points to yourself, what do you feel you deserve? Why?	

a 5. Obviously, something is amiss here. When this happens, I set the odd number aside and use the other scores for the average.

- ♦ About one group in five initially will have problems because one or two people are not participating adequately or are habitually late or absent. These problems can be corrected.
- ♦ It is essential that you give a practice peer evaluation about one-third or one-half of the way through the semester. The students fill these out and you tally them and give the students their average scores. You must carefully remind everyone what these numbers mean, and if

they don't like the results, they must do something to improve their scores. I tell them that it is no use blaming their group members for their perceptions. They must fix things, perhaps by talking to the group and asking how to compensate for their previous weakness. Also, I will always speak privately to any student who is in danger. These practice evaluations almost always significantly improve the group performance. Tardiness virtually stops and attendance is at least 95 percent.

When I look back at the essential methods of case evaluation, I am

struck by the fact that they are a mixture of objective and subjective elements. Some things are readily quantifiable and others are a struggle. Just as in courts of law where there are conflicting personalities, evidence, and judgment calls, teachers carry their own prejudices and experiences into the classroom. The image of a blindfolded justice rendering verdicts in an unbiased, impartial way has always struck me as unrealistic and, in fact, impossible. Justice, to be perfectly honest, must peek now and then and see who the participants are and what their histories are before making a judgment. So should teachers.