On Teaching and Learning

Editors:
JANINE BEMPECHAT
DEAN K. WHITLA

On Teaching and Learning publishes articles and essays on aspects of pedagogical practice and on research that has implications for teaching.

Letters and suggestions for articles or essays should be addressed to the Editors, 317 Science Center, Harvard University, Cambridge, MA 02138.

Published by the Harvard-Danforth Center for Teaching and Learning.
James D. Wilkinson, Director.

Faculty Development Writ Large: A Decade of Working with Teachers to Help Them Improve Their Teaching
Dean K. Whitla

The "Muddiest Point in the Lecture" as a Feedback Device
Frederick Mosteller

How Students Learn: Part II
Robert Rosenthal

Teaching by the Case Method: One Teacher's Beginnings
Nona Lyons

College Teaching is a Funny Business
Ian W. Brown
The "Muddiest Point in the Lecture" as a Feedback Device

Frederick Mosteller

Overview. To strengthen the teaching in a beginning course in statistics oriented to data analysis, I tried, with the cooperation of the students, an informal feedback system. Their suggestions led to some changes in the system, culminating in a three-step sequence:

Step A. In the final few (two to four) minutes of each lecture the students wrote brief responses to three questions: 1) What was the most important point in the lecture? 2) What was the muddiest point? 3) What would you like to hear more about?

Step B. I collected the responses on paper provided by the students (some signed, some not). For the next class meeting, I prepared a handout that gave a frequency distribution of the answers to the questions and that cleared up some inquiries or requests.

Step C. I tried to respond to as many of the requests as I could, some in later lectures as would have occurred naturally, some with extra handouts especially prepared, some with oral remarks in class, in addition to the remarks on the response lists.

The students were most cooperative, and some expressed considerable satisfaction with the method.

The setting. For readers not acquainted with beginning statistics courses, the content of this course is a peculiar mixture of mathematics, data, methods, ideas, and practical affairs, all related to variability—uncertainty—and, to some extent, to probabilities. Students usually take these courses to learn skills they can use later in their research or in their reading. Students come from life, physical, natural, or social sciences, or the humanities. For example, the course I describe had archaeology, law, biology, psychology, policy, political science, and business administration students among others. Inevitably the mathematical preparation of the students is likely to be uneven as is the number of years since they have last dealt with an x in cold blood. (Occasionally students are literally frightened of the mathematics.) A key problem is that a certain amount of detailed technical arithmetic is required, and a good many formulas of a fairly forbidding nature appear in areas such as analysis of variance and regression. In addition, the language often uses everyday words like confidence, significance, and power with very special technical meanings. Nearly everything is new. Lurking behind all this, students taking their first course in probability or statistics are forced to think about variability and uncertainty in every lesson. Many of the best students become more and more uneasy as the course progresses because they feel their deterministic world quivering and perhaps vanishing while they work, and most of them are unprepared for this. "If everything is so uncertain, how can I be sure?" I have taught such courses to many students at Harvard and elsewhere over many years. The course being described met three times a week for lectures plus a one-hour section meeting.

President Derek Bok encouraged the formation of a seminar on improving and evaluating college and university education. Under the leadership of Professor Richard Light, a master teacher himself, seminar members drawn from the faculties and administrations of many New England colleges and universities began meeting once a month at Harvard in the fall of 1986. Some presentations dealt with ways to get feedback from students so as to improve the course, usually when it is offered in the next semester or the next year. Toward the close of one talk, Patricia Cross told us of Richard C. Wilson's report (1986) of a remark of a physics professor with a fine teaching reputation. The teacher said that he wanted to know only two things at the end of a lecture: What the student thought the key point was and what the student wanted to know more about. He planned to ask this four or five times a course.

This simple idea attracted me because it might feed into this course given now and give immediate benefit to this teacher and these students without the need to wait for next year. This idea arose about the middle of the semester (total of 36 lectures).

The development. I implemented the idea gradually. I explained to the class about the President's seminar and requested that we begin with the first question, and they seemed willing. And so at the end of the hour they answered the question about the key point—some signed and some did not.

Some years ago Elizabeth Allison carried out a sequence of studies on teaching in the Department of Economics at Harvard. She found that the lecturer should summarize the key points in the lecture. I recog-
nized that not only didn't I always do that, but that in the frantic scramble of last-minute questions at the close of a lecture I wasn't likely to. Therefore, as a result of Allison's work, at the beginning of every lecture I write the key points, usually three to six, on the left-hand blackboard, where they sit all through the lecture. Since I began that, students have often mentioned this practice with approval in their end-of-course critiques.

Thus when current students wrote their single key point, they could have used one of mine, and no doubt some did. Many formulated the point in their own words. I summarized these. If a point was essentially one of mine, I recorded it that way; but if the student had a new formulation, I recorded it more or less in the words provided (taking some liberties because the student had little time to compose). Then I made a frequency distribution and handed it out at the next class meeting. It wasn't much trouble and gave few surprises. Of about 50 students, half to two-thirds responded; early winter always has absences, accounting for some of the non-response. As an illustration, List 1 gives the responses from a lecture asking only for the main point of the lecture.

<table>
<thead>
<tr>
<th>List 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>At the close of a class on multiple regression (MR) the class responded to the question “What was the most important idea in today’s lecture?”</td>
</tr>
<tr>
<td>No. of Responses</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

With this start I became braver and added the second question — What would you like to know more about? Again I got the answers and summarized and distributed them. After a couple of rounds of this, my view was that we weren't getting much payoff for the effort. The answer to what a student wanted to know more about was usually just the student's key point. List 2 illustrates responses to the two questions.

<table>
<thead>
<tr>
<th>List 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses to key idea and what you want to know more about.</td>
</tr>
<tr>
<td>No. of Responses</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
The "Muddiest Point in the Lecture" as a Feedback Device

WANT TO KNOW MORE

5 How to use the computer to apply these methods.
4 More on pros and cons of curvilinear regression.
3 More practical applications (including an example where line isn't better than the curve).
1 How you complete a stepwise regression (stop rules).
1 Why you adjust for the removed variables.
1 Transformation to straight lines (flat planes?).
1 See complicated function fitted.
1 More about R².
1 What has degrees of freedom to do with regression?
1 Responses the class offers to the two questions at end of the hour.

What the Lecture Addressed

The simplest methods of regression do their work by fitting lines and planes both because these are the simplest to fit, and because they often do the job, even though one might expect much more complicated methods to be required. Nevertheless, the methods used to fit lines and planes for forecasting can also be used to fit curves and curved surfaces. The lecture dealt with some of the pros and cons of fitting curves instead of straight lines. The simplest forms of curves add quadratic terms to the forecasts, and these quadratic terms tend to have very unstable coefficients. Consequently sometimes it is better not to use them for forecasting (as opposed to providing a scientific theory). Another main point of the lecture dealt with ways of choosing a good subset of variables for forecasting when there are many available — automatic science. Although computer programs can carry these out by methods called step-wise regression, investigators need to know that the programs are not magical and they need to keep control of the process.

The three students who ask for the responses the class submitted seem not to have grasped that we have been returning the responses at the next lecture every time. Perhaps Thanksgiving vacation led to some confusion. Or perhaps they are showing interest in the responses of others.

At this point I asked the students orally what they thought of the exercise: Was it worth it? Should we go on? They said that they liked the idea and pointed out that it gave them a quick opportunity to integrate the lecture at the end of the hour, and some regarded this as beneficial. They considerably pointed out that it was taking a good bit more of my time as well as secretarial time.

One student made a very perceptive breakthrough. He said, essentially, "The trouble with the procedure is that you aren't getting any feedback." How could he say this in the midst of what seemed to be massive feedback? He went on, "You give us the key points, and the ones we give you are much the same; and so you don't learn much about what we need, and our answers to what we want more of are too vague and general." Thus he had his finger on what I regarded as the blandness of the results.

Thus stimulated by him, I suggested that we might add the question "What was the muddiest point in the lecture?" This was met with acclamation, and we tried it that very day. This question produces a fountain of specific responses and leads to very concrete and relevant statements of what students want to know more about. Sometimes they ask for small things like how to carry out a calculation, sometimes for big ideas such as connections with other parts of the course, or even the outside world, and on a lucky day a few students don't have a muddy point. List 3 illustrates responses to the three questions for one day. We used the three questions for the rest of the semester except for one day when the class ran overtime.

<table>
<thead>
<tr>
<th>List 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses to questions about key point, the muddiest point, and what you want to know more about.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of Responses</th>
<th>KEY POINT</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>How to calculate BSS &amp; WSS &amp; F</td>
</tr>
<tr>
<td>10</td>
<td>TSS = BSS + WSS</td>
</tr>
<tr>
<td>1</td>
<td>Analysis of variance formulas</td>
</tr>
</tbody>
</table>
MUDDY

1 Unbelievable: Some people thought lecture very clear.
2 Calculating d.f.
3 Interpret within versus between.
4 ANOVA table.
5 Why do we use F-test?
6 Why do deviation scores add to zero within each group? [Because from the definition of the mean of a set of numbers, the sum of the deviations must add to zero. As follows:
\[ \frac{\sum x_i}{n} = \bar{x} \]
so multiply through by n
\[ \sum x_i = n\bar{x} \]
so for, say, n = 3
\[ x_1 + x_2 + x_3 = 3\bar{x} \]
Move 3\bar{x} to left side and break into \(- \bar{x} - \bar{x} - \bar{x}\) and get
\[ (x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) = 0; \]
SUM OF DEVIATIONS EQUAL ZERO, Q.E.D.]

10 Main point of F test (compare t and F).
4 Large practical example with detailed calculations.
3 More on interpretation.
1 Strength of evidence.
1 An example with a real effect. [Pulmonary example.]
1 Mathematical theory relating ANOVA to earlier part of course.

1 Why must we break TSS into BSS & WSS?
1 Where does F table come from?
1 [To a secret admirer — chocolate fudge.]
1 What about 2-way as well as 1-way tables? [Wednesday, I hope.]
1 Minitab statistics REGRESS command.

What the Lecture Addressed

The lecture was about analysis of variance. How do you allocate and measure the impact of sources of variation in experiments, such as effect on lung performance of various levels of smoking behavior, or the effect on the growth of pigs of different sorts of diets such as different kinds of proteins or the addition of a certain vitamin. Such variation is usually described ungrammatically as Between Groups (rather than Among), and Within Groups. Sums of squares of deviations from means play an important role in this allocation (a generalization of the Pythagorean theorem, where the square of the hypotenuse is equal to the sum of squares of the other two sides of a right triangle). BSS and WSS are abbreviations for between and within sums of squares. The F-test is a way of trying to find out whether group means vary more than can be explained by random sampling fluctuations based on within group variation. Note that some questions are oriented to the current lecture but others such as the one about deviation scores adding to zero were covered a month earlier, but two students suddenly want to know about it. This idea is important and worth shoring up because it comes up in so many circumstances. The lecture opened with some large examples of considerable stature and discussed the findings without yet teaching the calculations, so naturally the students are eager to find out how to carry them out in case I’m sneaky enough to put such things on the final exam, and I am.

ANOVA is an abbreviation for analysis of variance, and d.f. is an abbreviation for degrees of freedom, a key idea in assessing the reliability of the estimates. The requests for “MORE” are somewhat surprising. Some students have recognized that a statistic we studied earlier—t—is related to today’s new statistic F, and they want to find out what the relation is. That will be pursued in the
next lecture (whether they ask about it or not). Again, they want to see a large example with detailed calculations. They have already seen a good-sized example but we didn't do the calculations in detail. We did one in more detail at the next lecture primarily because of the request. We do much more on interpretation and detail. We do one in more detail at the next lecture primarily because we didn't try it later.

The request for an example with a real effect puzzled me because we have already had an example relating amount of smoking to lung performance with a slam-bang effect. Thus the response [pulmonary example] to indicate that they have had one. What I think I later found out through the "minute" papers was that "real" did not refer to the size of the effect but to the kind of example some students wanted to see. To them a "real life example" was not something with stature out of science, engineering, physics, medicine, or, say, public policy, as it was to me. Instead it was something sufficiently everyday that the student could personally relate to all parts of it, for example, something related to ways of learning and grades on standardized tests. For those students the pulmonary example was not satisfying because they didn't have any feeling about measures of lung performance, though they did about amount of smoking.

The final three questions listed show the variety of curiosity students actually have. In a non-mathematical course, it's a major effort to explain where the F-table comes from, and "Trust me" isn't much of an answer. I gave it a ten-minute effort at the next session, but to deliver the mail probably would require a substantial lecture and heavy homework which would benefit only a few students at this stage. The student wanting 2-way analysis (two sources of variation — say use of vitamins and use of different proteins in a diet) is thinking ahead. We did it later.

The story about the chocolate fudge answer is too long and pointless to explain here.

Finally the student worrying about the REGRESS command is trying to finish a computer assignment already overdue. He got help.

The main point of this list is that students are clearly concerned about a lot of different things, most of them connected with the current lesson. Some concerns can be met at the later lectures, others not within the confines of this course.

I have the impression that the students appreciate the effort, and a few have asked, almost with indignation, "Why haven't we been doing this all along?" The answer is that I didn't have this idea of instant feedback. I recognized the value of detailed written appraisals given by students at the close of a course in order to reorganize the course for the next time, but I had thought of the usual questions and answers during the lecture as handling the instant feedback.

**Conclusion.** My current thinking is that if the key points are already provided in writing, the question about the key point is not very helpful. But many kinds of lectures may not provide such key points, and then the question would still be profitable. I would not use it again as long as I offer the key points on the board or in a handout.

Some version of the "muddiest point" question seems valuable because it promotes concrete and sometimes non-trivial responses to the question of what "you want to know more about."

Things some students want to know more about will not interest others, or may require preparation others do not have. For example, some students wanted a mathematical derivation of a result we used. If I couldn't provide it in the page of responses, I prepared a handout. In all I prepared six handouts that I would probably not otherwise have prepared along with two more that I probably would have prepared anyway, (I usually give extra handouts, such as a least-squares approach to multiple regression, in response to student discussion, so this practice of handouts is not new, but these particular handouts were.) I have no notion of how this approach would work with a less sophisticated and less dedicated group of students. I think it will work for an instructor who is well armed against criticism.

I have been asked by the referee to say how much my behavior changed as a result of this approach. First, it cost me between half and three-quarters of an hour extra to prepare each summary three times a week. Probably about six minutes of the next lecture was changed from what it would ordinarily have been in an attempt to respond to queries. Then there was the time taken for writing up the extra handouts, probably an hour and a half each, on average. In addition, we stopped a couple of minutes early. All told then, perhaps 15% of class time (53 minute lectures) was changed, and we do not know who read the handouts.

I have also been asked about changed behavior of the students. They had to write up the "minute" paper, say three minutes, and many seemed to use that time to leaf through their notes to consider what to put down. A few who seemed not to have a class in the next period pondered their
responses for exasperatingly long periods while I waited, and they seemed very satisfied with what they had written. Those I talked with about it seemed pleased. I did not have the impression that anyone prepared a response before coming. Nobody complained to me about lecture time lost. Thus I would say roughly that their 5% changed 15% of my class behavior. And, of course, if you believe that participation speeds learning as most people do, this raised the level and maintained it at each session.

Discussion. Wilson says that the physics professor who suggested this feedback approach calls these responses by students “minute” papers. His students reported that he always knew if the class understood him. Wilson lists some other benefits from using “minute” papers (p. 199):

- It requires more active listening from students.
- It helps in identifying students who need special help or who may lack adequate preparation for the course.
- Our informant also tells us that it improves students’ writing. Responses during the last weeks of a class are longer and more articulate than those during the early weeks.
- It also helps document for students that they are indeed learning something in the course.”

Patricia Cross has suggested that “minute” papers could be used to gather many other sorts of information.

A minor drawback of this approach is that it invites only negative criticism, and I regard this as bad training. In my research work with students, which regularly involves criticizing the work of others, I emphasize that the first step must be to say something positive and concrete about the work. In inviting criticism, one is not asking to be blown out of the water, but for appreciation and help. In giving these short comments at the close of class, the student has no time for flowers even if trained to give them. I see no way around this except possibly to replace “most important point” by “best point in lecture,” which seems strained.

The method offers an attractive inexpensive way to get immediate feedback from the students in addition to classroom questions and answers. It supplements the great immediate feedback of silence. When nobody has any questions and the room becomes still, I assume that no one is following, and so I try to ask some easy questions to see where everyone got lost. The “minute” paper helps the teacher produce prepared responses at the next class meeting.

ACKNOWLEDGMENT

This paper was facilitated in part by Grant SES 8401422 from the National Science Foundation to Harvard University. I am grateful for comments and suggestions from David Hoaglin, Dean Whitla, and Cleo Youtz. The handouts of responses, special handouts, and the manuscript of this paper were ably produced by Marjorie Olson.

REFERENCE

Parting Thoughts

I have touched on but a small part of teaching here. The mechanics as to what to do and what not to do are obviously based on personal experiences and are geared to my personality. I do feel, however, that knowing the options may be of some value to you in your work. The thing which I have failed to communicate is how little of all of this is actually work. As I look back through this paper, I am saddened to discover that I have not really revealed the joy of college teaching. I was trained to do research, trained to increase knowledge, but it means so little if I do not have the opportunity to communicate it. How often I have wanted to shake a student by the shoulders and blurt out the answer for what appears to me to be so obvious. That is the temptation of the teacher. But how much more satisfaction I have received when a student has slowly plodded through the reasoning process and discovered the solution for him or herself. Teaching is indeed a funny business. How often you seem to be wasting time when you could just as easily do the job yourself. But as the years pass by and you see those very same students accomplishing things that you never dreamed of doing, you know you must be in the right business.

Acknowledgments

A more informal version of this essay was presented at the 1987 Fall Teaching Orientation of the Harvard-Danforth Center for Teaching and Learning. I would like to thank Eric Kristensen for inviting me to participate in this session and, thus, providing the stimulus for committing to paper what has been in my head.

The Contributors

Ian W. Brown is Associate Curator of North American Collections in the Peabody Museum of Archaeology and Ethnology. Mr. Brown teaches North American pre-history and a seminar on Indians in Museums at Harvard College. His research is on the archaeology of the lower Mississippi valley.

Nona Lyons is Lecturer on Education and Research Associate at the Harvard Graduate School of Education. A Spencer Fellow, Ms. Lyons teaches a course on teaching English, Social Studies, and the Humanities in secondary schools.

Frederick Mosteller is Professor of Mathematical Statistics, Emeritus at Harvard University. A gifted teacher, Mr. Mosteller has trained a generation of statisticians and has had an enormous impact on the study of statistics.

Robert Rosenthal is Professor of Social Psychology at Harvard University. A gifted teacher, Mr. Rosenthal teaches courses on research methods, statistics, and social psychology.

Dean K. Whitla is Director of the Office of Instructional Research and Evaluation in the Harvard Faculty of Arts and Sciences, and is Lecturer on Education in the Harvard Graduate School of Education. Mr. Whitla founded and directed the Harvard-Danforth Center for Teaching and Learning for the first decade of its existence.