Using Computer Spreadsheets to Assess Teaching Performance and Testing Effectiveness

Craig T. Evers
Purdue University - School of Technology

ABSTRACT

The traditional grade book can tell you how the students are doing in a given class. With the advent of computers and spreadsheet software, it is possible for teachers to also grade the tests and their own performance. Using Microsoft Excel, a spreadsheet has been constructed from which indicators of the testing effectiveness can be extracted. It is also possible to determine how well the teacher is doing in presenting class material. A series of interconnected tables plots the performance of each question in a test. By this means, with less effort than maintaining a written grade book, it is possible to evaluate the degree of comprehension the class has regarding any area of the subject matter being tested. It is also a straightforward matter to determine which questions have "worked" on a test and which have failed to evaluate the level of knowledge of the students. The development of a grading curve can be shown with every graded assignment so that the teacher always knows the degree of challenge to program into the next test or assignment. By apprising students of their standing at key points in the quarter, misunderstandings of grades and "surprises" are greatly reduced, and the students understand what is expected of them to reach their course goals. This tool was developed over a three year period of teaching engineering courses at Auburn University.

INTRODUCTION

A search of the literature and presentations at education conferences turns up many references concerning teaching and learning styles. As educators, these are vital areas in which we must be trained. There is much of value for us in these subject areas. Our students may not appreciate our preparedness, but if they learn, we will have succeeded in our charge to educate them.

In the proceedings of the 1996 ASEE Conference, for instance, there were papers dealing with a variety of topics, such as new faculty orientations[1], with the attendant training on teaching styles. Some schools have started to conduct seminars in college teaching[2], the better to prepare research-oriented graduate students for the realities of working with students. There was discussion of teaching styles[3], with reference to teachers who are at the baseline level, or who are active learning or discovery/design teachers. Bloom's taxonomy was invoked as a standard measure for learning and teaching. Tips, techniques, and tricks of the trade were shared freely on every hand.

There was one important area, however, in which there was very little information presented. There was a discussion of the teacher's perception of student academic success[4] in which we learned the extent to which our impression of a student's achievements varies from the students' self-expectations. Papers acknowledged that "student performance must be weighed," but passed

on quickly to deal with other aspects of the teaching experience. Some papers mentioned the use of different techniques to award grades, and there was talk of the “curve," but it seems that very little resource is focused on the development of this curve.

In the course of a lengthy industrial career, this engineer has participated in, taught, endured, and as a true student, slept through many training sessions regarding the use of metrics to determine project status. Industry uses many measurements to define success. These measures tell where we are along a planned course. They point out the direction we are headed, and give us information regarding areas in which corrective action should be taken to prevent falling below a nominal curve into an area where failure is a strong possibility. They allow a department or company to determine how they are doing and also to report to their customers whether expectations will be met. As educators, we have many metrics we use to define our success or failure. We must be skilled in their use to remain competitive in our own industry. Our customers, however, care about only one. The most important measure of success we use in reporting to our students is the grade.

Performance indicators can be tricky to apply properly. To be effective, a good metric must be versatile, powerful, and reasonably easy to apply. The chances are that if it is a bother to use, users will not bother with it. In teaching, we need metrics to measure the achievement levels of the class. A successful indicator should be able to tell us about the performance of the class as a whole. It should also give us feedback into our performance in teaching the class. Were there trouble spots in our coverage of a particular topic, areas in which class comprehension was weak? This is of special concern to new faculty teaching a course for the first time. How are the tests performing? Were there “bad” questions, or areas of poor achievement? A good metric should easily allow us to determine how we are doing as well as our students. Without such information, it is difficult to continuously improve our teaching abilities.

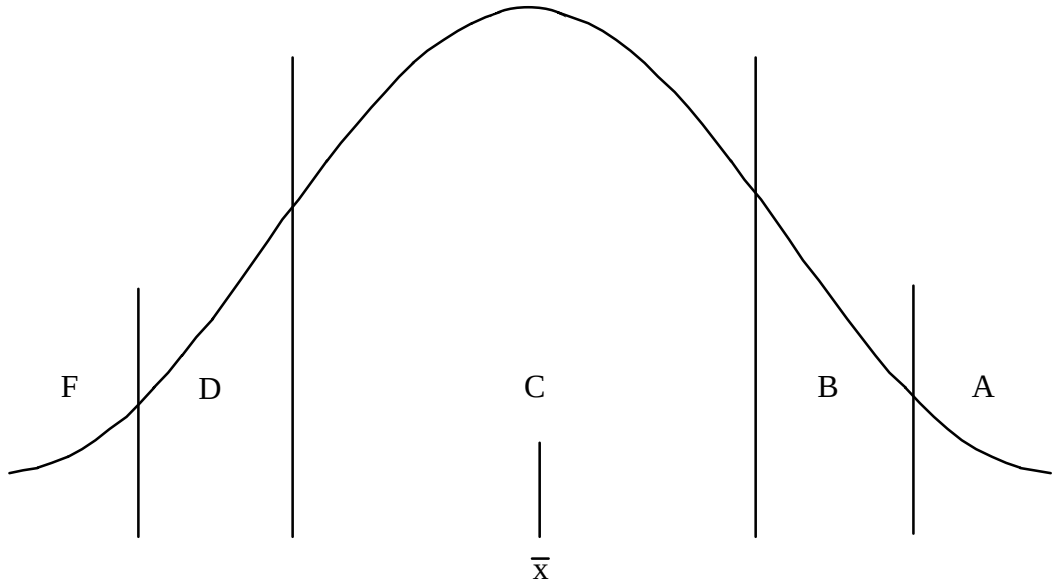

Figure 1 -- Normal curve theory of grades distribution

Statistical research tells us of a wonderful device called a “normal curve”. We often speak of this curve to our students. The words “I grade on the curve” can strike terror into the hearts of even the bravest undergraduate hopefuls. As shown in Figure 1, the curve represents the knowledge distribution found in a typical class of students. Some students will gain more than others. The average students will gather in the middle under the highest portion of the curve. As with most statistical concepts, a large class will give more validity to this idea than a small class, but the general idea is sound. There is a curve. The trick for teachers comes in matching that curve of knowledge with a predetermined numerical position on the scale of potential grades. How can we tell that a 90% cutoff actually corresponds to the top portion of the knowledge curve? Computer technology can provide us with a way.

METHOD

This grading spreadsheet was developed over a period of three years while teaching Industrial and Civil Engineering courses at Auburn University. It has been further refined in Mechanical Engineering Technology courses over the last year at Purdue University’s Elkhart Technology site. The original motivation for the tool was that computer resources were readily available, while a special gradebook was not. Over the course of time, its role grew from one of recording grades to a diagnostic teaching evaluation tool.

In its earliest form, the chart had rows for each students’ name and columns for the numerical grade on each assignment and test. As experience was gained with this technique, it was found to be convenient to build a separate table for each test. This allowed simple calculation of the grade for the test, and easy transfer of data into the correct column of the main spreadsheet. It was then a short conceptual step to analyze available data in more detail. Charts of class performance were found to be a valuable aid in helping students understand their exact standing in a class.

As the power of the tool developed, the next step was to analyze tests for overall and specific performance. As shown in Figure 2, it is a simple matter to pick out questions that did not contribute to the development of a curve of knowledge. Questions in which the entire class did poorly indicate either that the class failed to comprehend a specific area of instruction or that the question was poorly constructed and the students did not understand what was sought. In either case, there may be opportunities for some discussion with students regarding the “fairness” of these questions and the way they were graded.

Using Figure 2, the instructor can analyze the test for success. This figure (populated with fictitious students and scores for illustration purposes) shows that as a whole, the class did fairly well on the test. The class average was 81.83, with a standard deviation of 11.65. One student did very well, one did very poorly. Closer examination reveals that one question is suspect. Question 2 appears to have not worked well in developing the curve of knowledge from this class, with only 55% of possible points being awarded. The teacher should review this question to determine the reason for the low scoring. If the test questions actually represent course objectives and established standards, and they do not perform well on the test, the scores may indicate that the material was not presented well in class.

MET 2EZ	Summer, 1997			Midterm Test Results			
Question #	1	2	3	4	5	6	Total
Points	15	10	25	20	10	20	100
Greeley, AB	15	4	20	19	10	19	87
Jackson, CD	14	10	23	20	10	20	97
Lewis, EF	10	6	13	12	5	16	62
Richardson, GH	12	3	19	16	10	19	79
Sill, IJ	12	6	21	19	8	20	86
Wood, KL	11	4	22	15	10	18	80
Earned	74	33	118	101	53	112	491
Possible	90	60	150	120	60	120	600
Ave	12.33	5.50	19.67	16.83	8.83	18.67	81.83
Std Dev	1.86	2.51	3.56	3.06	2.04	1.51	11.65
High	15	10	23	20	10	20	97
Low	10	3	13	12	5	16	62
Percent	82.2%	55.0%	78.7%	84.2%	88.3%	93.3%	81.8%

Figure 2 -- Sample of Midterm Test Results

In a slightly modified format, Figure 2 also makes an excellent tool for the students to use in self evaluation. With the names removed and the rows of students sorted by total score, the students can find their score on the table and see exactly how they did relative to their peers.

The “total scores” column can easily be copied from this table and the values pasted into the master spreadsheet shown in Figure 3. In this form, the real power of this tool is realized. The teacher can easily plot curves of class performance. From these, break points can be determined for grading. Offset margins or “curves” can be tested for their effect if it is considered appropriate to alter the scores.

As much as teachers hate to admit it, poor performance on a test is not always the fault of poor students. There are occasions when a teacher will fail to either teach to the degree of difficulty expected in the test, or will simply “blow it.” If this happens, it is easy to spot in the spreadsheet. A bimodal distribution of grades indicates that some students understood what was taught, and others did not. This is an opportunity for creative introspection as the teacher develops recovery techniques. Having the data to quantify this situation helps in dealing with it.

RESULTS

When a student comes in for help, it is a simple matter to review the spreadsheet and find areas where extra effort can make a difference. In Figure 3, EF Lewis appears to be trying. He has turned in all assignments and taken all quizzes. His performance in these areas and his midterm score seem to indicate not so much a lack of effort as a lack of comprehension. If he had not been turning in homework, it would be easy to spot. This may help in counseling. The average he has achieved is clearly identifiable and can be compared to the class average for each assignment or

area. In this respect, the spreadsheet functions as a traditional grade book, but with much greater ease of updating calculations.

MET 2EZ - Grades Summer, 1997

Name	Homework Scores Ch. 1	Ch. 2	Ch. 3	HW Total	Quizzes 1	2	Quizzes Total	Midterm	Final	Grand Total
Possible	15	15	25	55	20	20	40	100	150	345
Assigned	15	15		30	20	20	40	100		170
Greeley, AB	11	15		26	15	20	35	87		148
Jackson, CD	12	12		24	20	20	40	97		161
Lewis, EF	8	9		17	12	8	20	62		99
Richardson, GH	14	15		29	13	10	23	79		131
Sill, IJ	10	12		22	17	10	27	86		135
Wood, KL	13	15		28	20	20	40	80		148
Average Grade	11.33	13.00	####	24.33	16.17	14.67	30.83	81.83	######	137.00
								11.65	######	21.46
								97	#NUM!	161
								62	#NUM!	99

Points to Date		Grade	Code
Greeley, AB	87.1%		
Jackson, CD	94.7%		
Lewis, EF	58.2%		
Richardson, GH	77.1%		
Sill, IJ	79.4%		
Wood, KL	87.1%		
Ave	80.6%		
Std Dev	12.6%		
High	94.7%		
Low	58.2%	Grades	

1	0	90 - 100%	1	0	A
3	0	80 - 90%	2	0	B
1	0	70 - 80%	2	0	C
1	0	60 - 70%	0	0	D
0	0	below 60%	1	0	F

Figure 3 -- Grading Spreadsheet

When returning tests, it has been found very helpful to write each student's total score to date on the test as well as the test score. In this way, students can add up the scores of all returned work and verify the accuracy of the teacher's record. If discrepancies are found, there is time to deal with them before the crush of finals week. This has greatly reduced last minute plea bargaining and appeals for mercy.

From the standpoint of student relations, an unintended effect of reporting grades in this manner is that the students have an understanding of how the teacher is tracking progress. With this understanding, the grading process becomes more personally relevant, and they accept responsibility for their performance and its recording. There is a greater understanding of where they stand and what the goal is for satisfactory accomplishment. With the spreadsheet showing test performance, students can identify and compare their results to the rest of the class in a non-threatening manner, without revealing their identities. The better students realize that these spreadsheets may also be a useful predictor of areas of emphasis in future tests. Broad hints from the teacher see to that.

In addition to helping the students, this tool is also a good diagnostic tool for the teacher to evaluate his or her performance. In a recent class being taught for the first time, the first midterm was, to put it bluntly, disastrous. The curve of knowledge revealed by the spreadsheet showed as many students in the 90's as were in the 30's, with few in between. Responsibility lay largely with the instructor's expectations. Methodology developed at one school did not translate well to another with a different student population and subject area. The grade curve was very low.

Changes in teaching style and methods of evaluation were made and the results monitored for the next test cycle. Improvements were clearly visible. By providing adequate opportunities to earn points, and monitoring the development of the class curve, a full recovery was made by the final exam, and grades were accepted as fair by students and teacher alike.

This method of correcting a mistake in test performance seems more defensible than announcing a blanket adjustment to all test scores. The problem with “curving” the grades is that students come to expect it as an antidote to less-than-satisfactory performance, and they lose faith in a method of grading which requires such adjustments. It is their future on the line, and they like to see a method they can trust at work in the process. With this tool, it is possible for the teacher to track the class standing. In Figure 3, there is a section in the bottom of the chart for histograms of class placement. The results of the midterm show a different curve than the total to date. The next assignment may be adjusted in its level of challenge to spread or tighten the curve of knowledge as the teacher thinks is needed. Whatever grading technique (normative, criteria, etc.) is used by a teacher, the spreadsheet gives the data necessary to implement it well.

When teaching a class, subtle changes may influence the degree of comprehension from one class to the next. This is especially likely with new instructors whose experience with the subject matter or teaching are changing rapidly with each iteration of the course. Problem areas are easily identified. Given a choice, most would think it a more effective use of the final to retest an area that scored lower than one that showed good comprehension. Specific questions can be revised for the final to give students another chance to explain their understanding.

DISCUSSION

In general, students are surprised at the amount of information this tool provides. Responses are uniformly positive once they recover from the shock and realize what they have been given. They become, in effect, part of the process, not just spectators hoping for a good grade at the end of the term. Students indicate a sense of control as they realize how their performance affects their standing and grade.

When a student has a progress question or concern, it is an easy matter to print out only the line of the large spreadsheet for that student, together with the column headings. The student then has a record against which to check his or her own records. For discussions on the phone, the teacher can review each area of grading, and advise the student how it compares to the class average or class highs and lows. In this setting it is easy to identify areas for improvement.

When scoring tests, a numerical value is entered in the test margins, and then all scores are entered into the spreadsheet (see Figure 2). Totals are read from the right-hand column and recorded on the test sheet. No calculators are needed to add up each page. No additional transcription of results is needed. The total column is pasted directly into the appropriate column of the Figure 3 spreadsheet, where the points to date and histograms are automatically updated. The curve of knowledge of the class is immediately available for review. Totals can be plotted in a chart if desired for visual determination of break points in the continuum of grades.

This grading tool has been found to be an extremely effective means of evaluating both the students and the teacher. Developed originally as a time and labor saving tool, it has exceeded these objectives. It has become an invaluable technique to analyze the results of much that takes place in the class, and helps identify problems with both learning and teaching. When properly applied, it reduces student concerns about the grading process, and helps the teacher to more accurately assess and adjust the knowledge curve of a class.

Biographical Sketch

CRAIG T. EVERS is an assistant professor in the Department of Mechanical Engineering Technology at Purdue University, teaching at the Elkhart, Indiana site. He has a background in manufacturing engineering spanning over 16 years with companies such as IBM, Kimball, and the Robert Bosch Corporation. He has a BSME (Utah State University, 1982) and an MIE (Auburn University, 1993). He is also a registered Professional Engineer.

REFERENCES

[1] Predebon, William W., Cho, Peck, Ott, Linda M., Sweany, Philip; New Faculty Orientation and Seminar Series: Emphasis on Teaching and Learning, ASEE Annual Conference Proceedings, 1996

[2] DiBiasio, David, Miller, Judith E., Groccia, James E.; Managers of the Learning Process: Preparing Future Faculty to Teach Productively, ASEE Annual Conference Proceedings, 1996.

[3] Samples, Jerry W., Nygren, Kip P.; Becoming a Better Teacher: Adjusting from the Baseline, ASEE Annual Conference Proceedings, 1996

[4] Bogales, HR, Shiber, JG, Teachers' Perception of Student Academic Success, ASEE Annual Conference Proceedings, 1996.