Session 1460

Breaking the Curve - Why a Straight-Scale Is Appropriate in Engineering Courses

Dr. Paul Blowers
Department of Chemical and Environmental Engineering
The University of Arizona

Abstract

Many instructors have a tendency to place students on a curved grading scale based upon statistics and the average student's performance during the semester. This is often done because it makes it easy for professors to assign final grades. Professors can also feel more comfortable about their exams because any unfair questions will be normalized out through the curve. It is comfortable to use a curved scale, too, because many of the courses the instructors had were taught on curved grading scales.

There are several problems with adopting a traditional curved grading scale within engineering courses. We will highlight several points that should demonstrate why a straight scale is more appropriate.

Curved grading scales are based upon statistical assumptions that the class is large and that the student body is randomly dispersed over a wide range of performance. However, most engineering courses do not have hundreds of students in them so the size of the class could skew the results. Another problem is that engineering does not attract a random sample of student abilities. Within the three universities that I have been affiliated with, each school had a predominantly higher percentage of honors students enter the engineering program than the general student population. In the case of the University of Arizona, 18% of students are in the honors program, but 48% of chemical engineers are. One should ask whether it is appropriate to curve the scale when most of the students are of very high ability. This would punish students for being bright enough to choose an engineering career. So, one should question how large the class is and the raw student caliber entering into the class before using a curved scale.

A larger problem with curved grading scales is that they undermine many of our other goals as educators. ABET accreditation requires that our students are well versed in communication and teamwork skills when they graduate. A curved grading scale inherently sets up competition between classmates because any aid they give to another student may raise the class average, making it harder for them to receive a high grade. This sense of competition may prevent students from forming true collaborations on their team projects and may prevent them from learning how to effectively communicate. For all of these reasons, we strongly advocate that all engineering courses should be taught on straight scales.
I. Introduction

"A grade is an inadequate report of an inaccurate judgment by a biased and variable judge of the extent to which a student has achieved an undefined level of mastery of an unknown proportion of an indefinite material." - Paul Dressel.

As educators, we must both be advocates for and judges of our students. This sets up an internal conflict because we would like our students to do well and encourage them to do their best. We must also evaluate them on their progress and often tell them when they are not performing to adequate standards. To minimize internal conflicts, faculty seek grading methodologies that will avoid these problems. The four major approaches to grading include: norm (grading on a curve), criterion (grading on a straight scale), mastery, and pass/fail. This paper will address the issue of whether engineering courses should be graded on a curve or should be graded on a straight scale.

Because we are engineers by training, faculty in engineering departments are always searching for efficient solutions to complex constrained problems. One of the complex optimizations we must do is to balance our time constraints with a complete and correct accounting of student learning within our courses. The time balance between all of our duties and the need to rate student learning has often led to the use of curved grading scales within our courses. Across the nation, as many as 30% to almost 50% of faculty in engineering departments are choosing to grade on a curve.

Grading students on a curve involves using statistical criteria for selecting where breakpoints in a grade distribution will fall. The average course grade and standard deviation are commonly used to find the break points, as shown in Figure 1. With spreadsheet software and easy data manipulation, finding a curved scale is easy to do. It is also attractive because it is based on mathematical arguments from statistics that seem to validate its use. It is also a comfortable tool to use because it is familiar to all college students and we grew accustomed to its use when we were in school. However, using a curved grading scale has no real mathematical basis. It just happens that some traits do fit the "Bell-Curve" in large populations, although learning has not been proven to be one of those traits.

Figure 1 - A commonly used curved grading scale. Here, the average score is the cut-off between a B and a C while one standard deviation segments are used for other cuts.
For an excellent review of the history of grading on a curve see Guskey's work.\(^6\)

Grading on a curved scale can also eliminate many difficulties that arise during the semester as well. Sometimes faculty are rushed while preparing an exam or complex assignment and just are not able to develop a full and fair learning assessment tool. If a question is unfair or is extremely difficult to solve, then the curve and the average will be lowered. This will effectively normalize out the difficulties with the question so that students who performed statistically better will receive a better grade while those who performed worse will receive a lower grade.\(^7\) In effect, using a curve ensures that any mistakes we make when constructing an assignment or test will be normalized out at the end of the semester when the final grade distribution is set.

II. Statistical Analyses Related to Using a Curved Scale

When we use curved grading scales, there are several assumptions that we hold as valid. First, a statistically based grading scale assumes that there is a large population to ensure the analysis will be meaningful. Second, a normal distribution of student performance with the infamous Bell Curve shape is assumed. Finally, other external factors which could shift the scale are ignored. We will now discuss each of these inputs to using a curved grading scale.

One must ask how large a class must be in order for a curved scale to be used in assigning final grades before one can be comfortable using one in a class. The Bell Curve has been found to be true of very large populations, normally in the thousands or millions. However, enrollment data for many engineering departments shows that the average class size is less than 40 students.\(^8\) While some larger schools may have large numbers of students and may begin to approximate the Bell Curve distribution, the bulk of engineering classes do not have that luxury.

Another of the underpinnings of the curved grading scale is that there is a normal distribution of student performance during the semester.\(^2\) This leads to the Bell Curve in Figure 1. However, what happens to the statistical meaningfulness of using a curved scale when the distribution of student performance is not normal? Figure 2 shows a typical bimodal distribution of student performance compared to the normal distribution.

Here, the mathematical underpinnings may lead to unusual consequences. A very high standard deviation may
make it impossible for any student in the class to receive an A even if they received 100 percent of the total points possible. This can happen when the distribution is far from the assumed bell curve shape.

A final effect commonly encountered in our engineering courses is that there are external factors outside our control which suggest that a curved grading scale is not appropriate. For instance, are our students truly "average students"? Are they the same as the rest of the student population so that we can feel comfortable using a curved grading scale? I will use some numerical data from the Department of Chemical and Environmental Engineering at the University of Arizona to highlight this point.

At the University of Arizona, approximately 18% of the general student population participates in the campus-wide honors program\(^9\). However, a closer look at the Department of Chemical Engineering reveals that 48% of chemical engineers are involved. There are many highly motivated and bright students who have chosen chemical engineering for the intellectual challenge they will receive in college and the high starting salary the degree provides upon graduation. So a larger percentage of very good students are choosing this engineering department over other all other departments within the same university. These comments are echoed by others across the country\(^{10-11}\) and we must ask what implications this shifting of raw student caliber will have on student assessment within the classroom.

![Figure 3 - Statistical numbers of honors students at the University of Arizona](image)

![Figure 4 - A comparison of a "normal" class with a high caliber class: unfairly penalizing the moderate performer.](image)

Figure 4 shows a comparison between a "normal" class and a class filled with higher caliber students. We see that when the class averages differ significantly, the grade distribution also shifts radically. A score of fifty percent in the "normal" class would receive a high C or low B while the same score in a class filled with higher caliber students would be a failing grade. In fact, an average grade in the "better" class would be an A in the "normal" class. The implication of shifting course grades like this is that we are choosing to penalize highly motivated and
bright students for choosing engineering when we use curved grading scales. We will now consider a hypothetical extreme case to make this point even further.

Imagine you are teaching an introductory physics course and you only have two students in your class. You decide that a curved grading scale is appropriate so one student will receive a C, D or F while the other student receives an A or B. Now imagine these two students are Albert Einstein and Stephen Hawking. Which student would fail and which one would not? In this hypothetical example we see that the size of the class and the caliber of the raw entering class can greatly affect how a curved grading scale can penalize students who are very good at what they do if they end up in a class of similar students.

We all have had year to year variability in the student quality entering our classes. One could argue that it is not fair to use a curved grading scale because students who happen to end up in a higher caliber class will do worse than if they were in a "normal" class. The students would not necessarily know any less material in the higher performing class, but their grade would reflect that they had in fact done worse than if they had been in a lower averaging class. On the contrary, we often see a class of highly motivated students that pull all of their classmates to a higher level of competence; all students learn more when the class caliber is statistically better than average years and when the students collaborate with one another.

At this point we have discussed some of the reasons why curved grading scales have been used to evaluate student performance. We have also discussed why a statistical averaging of student performance may fail due to small class sizes or unusual grade distributions among students. We also put forth the claim that our students are not average students from a general population and that they should not be penalized for choosing a rigorous and rewarding engineering degree program. Now we will turn to the broader implications of using a curved grading scale in engineering courses.

III. ABET Criteria and a Curved Grading Scale

We are all familiar with the new ABET criteria where we must continually strive to improve student performance and our inclusion of several traditionally non-engineering skills. Two newer areas where are students must demonstrate proficiency are:

- an ability to communicate effectively,
- and, an ability to function on multi-disciplinary teams.

Grading on a curve makes it much more difficult to accomplish these goals on a class-wide basis.

Some students may resist communicating with their classmates and working in teams if one is grading on a curve. The student motivations that affect the success of the ABET criteria are shown in Figure 5. Students near the top of the curve may say, "I'm not going to help anyone else in the class because that will just raise the curve higher. I'm certainly not going to help anyone that is doing well in the class because I want the top grade so I can guarantee an A." Here, the student feels that they are in competition for a limited number of A grades that may be distributed and they want to ensure that they get one of them.
Students near the middle of the curve will feel similarly to those near the top of the curve. They may say, "Talking with other students and helping them learn will only make it harder for me to get a good grade in this class. It's hard enough already!" We again see that students will not want to communicate with others because sharing information and working together will just shift the curve higher and make it more difficult for students to do well against the curve.

Students near the bottom of the curve may have other difficulties. These students may truly desire to work with others and to ask them questions. However, after the other students in the class have refused to communicate with them a few times, they too will give up and no longer communicate with their classmates.

Grading on a curve makes it more difficult to encourage students to develop their communications skills and to work in teams. In some instances, the sense of competition can be a motivating factor and can lead to increases in learning. However, one could make the argument that using a rigorous straight grading scale can also provide a challenging environment where students are encouraged to collaborate on complex problem solutions. This leads to the students competing against the material instead of each other, which is desirable because the students' true goal should be mastery of the material and not a high grade in the course.

It is much easier to sell ABET requirements to students when they are not penalized for trying to fulfill them. The communications and teamwork criteria are particularly difficult to get students to buy into when a curved grading scale is used. A complete review of other assessment issues regarding a diverse range of educational impacts is available.
IV. Changes in Student Motivation for Other Reasons with a Curved Grading Scale

Both high performing and low performing students will often change how they allocate their efforts among their classes when a curved grading scale is used in one or several of them. Students who normally do very well in their courses compared to their peers may choose to allocate less effort in a course when it is graded on a curve. Their rationalization would say, "I normally get an A against these same students anyway. So, I won't work as hard in this class because it is graded on a curve. Instead I'll work on other classes or do something else and I'll still probably get an A." In this case, a student who normally would be very motivated to learn the material has decided the grade is important instead of their actual increase in knowledge after they complete the course. The student will actually learn less in this situation than if the course were on a straight scale.

On the other hand, underperforming students may use a different rationalization when a curved grading scale is used. Here, students may say, "It doesn't matter how hard I work against my classmates because I never do well compared to them. I'm not even going to bother trying very hard in this class because I know I won't be able to get an A against them. I just want to pass." The student again may reduce their effort on the curved class because they feel that investing time in it is not going to improve their grade. Effort would be wasted, they feel, because they have no hope of doing well.

Also, not all cultures are based on competition. This means that students coming from non-competitive cultures may feel uncomfortable or be unwilling to participate in class activities. It is unfair to place these students on a curved grading scale\textsuperscript{13}.

Grading on a curve also creates a very difficult decision scenario for students when a project, grade, or exam can be dropped at the student's discretion. In this case, students are left to decide which grades to count and which ones to eliminate. Students will normally have the class averages on the exams and their own scores when selecting which grade to remove. Should the student drop their lowest absolute score? Or should the student drop their grade on the assignment where they had the lowest score compared to the average?

Mathematics would say that the student should drop their lowest absolute score since it is contributing the least value to their overall grade, but students may not think this way. As students select which scores will count and which one will not, the class average will shift. If students knew beforehand how scores would be changed and what the new averages would be, they might choose a new score to drop. This process becomes a psychological and mathematical morass to the students as they try to perform a very ill-defined non-linear optimization. Students just do not know where they clearly stand in the class until the final grades are distributed in this kind of scenario.

V. Conclusions

Grading on a curve may seem like an attractive way to efficiently assess student performance in engineering classes. However, grading on a curve can penalize bright students because they have chosen engineering along with other bright classmates. Course sizes may not
be large enough and performance distributions may not be normal for statistically based methods to be valid. Students on curved grading scales are generally motivated by grade based considerations to learn material instead of an innate desire to learn. They view their classmates as competitors that must be defeated instead of as resources to help them learn. Communication among the class members and the ability to create effective teams, both ABET criteria, are hurt by this competition among class members. For all of these reasons, curved grading scales should not be used in engineering courses.

References

7. Wankat, P., and F. Oreovicz, "Knowing your way around a curve will make grading less arbitrary, and earn you high scores from your students", *ASEE Prism*, 8, 18, (1999).
8. An estimate based on data from Figure 3-3, page 29 of *Engineering Education: Designing an Adaptive System Board on Engineering Education*, National Research Council (1995). Also based on a random survey of engineering class sizes reported on college websites.

PAUL BLOWERS

Paul Blowers is an Assistant Professor in the Department of Chemical and Environmental Engineering at the University of Arizona. He received his B.S. in Chemical Engineering from Michigan State University before attending the University of Illinois at Urbana-Champaign for his M.S. and Ph.D. in Chemical Engineering. In addition to educational research, his other academic research involves using quantum chemical techniques for predicting reaction rates in different environments.