Modeling Undergraduate Student Effort: Exploring the Gap Between Effort and Grade

Jannett Highfill
Bradley University

Tanya M. Marcum
Bradley University

While excellent reviews of student behavior relative to study time are available (e.g. Oettinger, 2002; Babcock and Marks, 2011; Allgood, Walstad, and Siefried, 2015), the contribution of the present paper is to argue that instructors may choose to adopt policies or practices that affect the gap between student effort and course grade and that such practices might usefully be characterized as impacting the average gap or the dispersion of results (the standard deviation). In addition, we include a simple model of student effort and reflections from instructors based on experiences and perceptions.

INTRODUCTION

Teacher-scholars, dedicated to thinking hard about teaching and learning, inevitably have explicit or implicit models of student behavior that raise the question of the relationship between student effort and grading. Babcock and Marks (2011) somewhat famously found that the average study time in 1961 for undergraduates was forty hours a week while by 2004 it had fallen to twenty-seven hours a week. It seems that study time has, for whatever reason, become scarcer which perhaps increases the challenges instructors face in assigning grades. The student effort literature typically assumes (see the excellent survey by Allgood, Walstad, and Siefried, 2015) that student utility is a function of learning or knowledge, grades, and leisure; the choice variables are the amounts of study time a student allocates to each of their courses. Subject to such things as instructor time, pedagogy or technology, and the grading policy in a given course, all of which are parameters from the student's point of view, the student chooses their study effort (study time) for each course to maximize their utility function.

The literature also recognizes that the relationship between study effort and the resulting grade is not as straightforward as put in the time, get out a grade. The focus of the present paper is on the potential gap between the two, study effort and grade. We will provide examples from our own experiences and from those of our colleagues in what follows, but to get started, imagine that an important case study in a particular course involves a university choosing the location of its law school that is currently located in the Loop District of Chicago. Suppose a student in class has a spouse that just graduated from that law school and loves it. Another student, on the other hand, had a spouse that had a bad experience there and they try as best they can to forget about it. The accidental details of the case might well imply that, other things being equal, the first student will have a leg up on the case compared to the second student because of differing emotional valences.

To help us think about these situations we will rely on a model adapted from Oettinger (2002) which assumes a student effort function in which the number of points a student receives and the ultimate grade earned are subject to a random component. We will use the model to investigate the gap between effort and grade using some simple numerical examples. To be clear, the paper is not empirical nor does it advance the theory in a substantial way. Rather the goal is to reflect on the implications of the gap between effort and reward for students, and then for instructors. Instructors may influence to some degree the factors that students must take as their parameters in their decision-making strategies in a particular course. This paper will argue that instructors may choose to adopt policies or practices that affect the gap between student effort and course grade and that such practices might usefully be characterized as impacting the average gap or the dispersion of results (the standard deviation). A concern with student effort and grading will also lead to some very brief thoughts on grading systems themselves, in particular whether grades should be the A, B, C, D, and F as is common at some universities, or whether a plus/minus system is preferable, or for that matter some other system such as "A/Pass/ Fail". Some institutions have slightly different versions of the "A-B-C-D-F" scale such as "A, AB, B, BC, C, CD, D, and F" (Perlman, McCann, & Prust, 2007). While no instructor chooses their university's convention on this question, at the most abstract level it is nevertheless a matter decided by the faculty.

BRIEF LITERATURE REVIEW

Higher education has recently felt the pressure to adapt teaching styles to reach all students and to motivate them (Forsyth, Kimble, Birch, Deel, and Brauer, 2016, p. 84). Instructional materials explaining the instructor's expectations for a particular course may assist the student with their scheduling and use of time as well as expectations for particular course grades (Hamer and O'Keefe, 2013, p. 27). The concept that study time affects grades has not always been supported by the empirical literature. But with the refinement of methods increasingly an effect has been found. Stinebrickner and Stinebrickner (2008) found that an increase in hours of study time per day had a positive effect on overall student GPA. Bonesronning and Opstad (2012) found a smaller effect, but it was for a specific course not overall GPA. Bonesronning and Opstad (2015) also found a positive relationship between study effort and learning achievement. Their strategy was to study a natural experiment in which a midterm was added to a course that did not previously have one. Students responded by changing in study effort. After the exam students that performed "better than expected" decreased their study effort while those that performed "poorer than expected" increased their study effort.

Somewhat reassuringly, this is evidence that factors under the control of the department or the instructor will affect outcomes. To mention just a few other examples, Ho and Kelman (2014) found that modifications in the grading system and reducing class size reduced a law school gender gap. Brustin and Chavkin (1997) and Rose (2011) respectively studied such issues as pass/fail grading and grading on the curve. Peterson and Peterson (2016) adopted a grading approach based on 10,000-points, wherein the scale for an A was 9,000 to 10,000, 8,000 to 8,999 would be a B, and so on. This approach was adopted due to student pressure to be graded based on effort. In the experiment by Brustin and Chavkin (1997) in a lawschool general clinic course, students were allowed to select either pass/fail or full-grades just prior to the drop date for the course.

We mentioned the decline in study effort on average reported by Babcock and Marks (2011). Grade inflation is the symmetrical concern. Butcher, McEwan, and Weerapana (2014), for example, examined policies used to reduce grade inflation, Hernandez-Julian and Looney (2016) explicitly connect the issue to grading practices, and Charlton and Stevens (2000) attempt to construct a methodology that could be used with registrar data to reveal the grading standards and rigor of the curriculum on a course-by-course basis.

Concerns with grade inflation suggest the question of whether a tougher grading standard induces more or less study effort. The positive effect might be because grades serve as motivation; the negative effect might be because students are discouraged by unattainable standards. Grant and Green (2013) in a provocative study of grades as incentives, argue that current research essentially concludes that "intrinsic" motivation and "extrinsic" motivation are substitutes not complements in that students have an "achievement motive" that is *weakened* by the use of incentives (2013, p. 1566, italics in the original). Empirically they find that the "indirect rewards provided by grades" do not motivate students to strategically study more when it is most likely to benefit them, i.e., when they are near but below a grade cutoff (2013, p. 1582). In another case Babcock (2010) found relatively higher study effort in a class with a relatively lower expected grade.

A SIMPLE MODEL OF STUDENT EFFORT

Suppose that students care only about their course grade and their study time with an eye to balancing their desire for a grade with time to pursue other interests or accomplish other goals. That is, suppose that the student's utility is only a function of the grade and leisure, and the only constraint the student faces is a time constraint, where any time not spent studying is "leisure time." Study effort and study time are equivalent. The student's performance, in terms of points earned over the course of the semester, is a function of study effort (assuming student ability is held constant). It is assumed that there are a finite number of grades with specified cutoffs for the grades. In a world of complete certainty, the student would allocate the study effort needed to get the minimum number of points required to achieve the desired final grade.

But suppose student effort is subject to a random shock that introduces a possible gap between study time and final course grade, with the shock being a continuous random variable which is (1) unimodal, (2) has a mean of zero, and (3) is symmetric in that positive and negative shocks of a given size are equally likely. In practice, following Oettinger (2002), the shock will be modeled using the normal distribution. These assumptions assure that the determinate outcome of study effort is an unbiased predictor of the course grade. (Please see the technical appendix for a formal statement of the model.)

For a simple example, suppose that the student's work effort in a world without uncertainty would have earned a grade of 87, and suppose that the bottom cutoff for an A (for the moment assuming no pluses or minuses) is a 90. In a certain world the student would get the next lower grade, say a B in the A-B-C-D-F grading system or a Pass in the A/Pass/Fail grading system. For simplicity we will consider the latter grading system first. The student would get an A if the random shock was such that they received an additional +3 or more points. Supposing that the bottom cutoff for a Pass grade was 60, the student would Pass unless the random shock was larger (more negative) than -27. (Assuming a continuous probability distribution but discrete number of points and cutoffs is sometimes slightly awkward; for example, the probability of exactly -27 is zero as is the probability that a student falls exactly on any grade cutoff. We make these assumptions following the lead of Oettinger (2002) but occasionally make a few adjustments in the verbal discussion of results.)

We will begin our numerical examples with the A/Pass/Fail grading system because of its extreme simplicity. For each table we consider three cases, in the first case the standard deviation is chosen so that there is a 95% probability that the random variable is between -1 and 1. This is called the "low standard deviation" case. The "moderate standard deviation" case is where there is a 95% probability that the random shock is between -2.5 and 2.5. The "high standard deviation" case assumes there is a 95% probability that the random shock is between -5 and 5. (All probabilities in the tables are rounded to one significant digit.)

TABLE 1
A, PASS, FAIL PROBABILITIES

Earned Grade	91	89	61	59
	Low Standard Deviation			
A $y \geq 90$	97.5%	2.5%	0.0%	0.0%
P $60 \leq y < 90$	2.5%	97.5%	97.5%	2.5%
F $y < 60$	0.0%	0.0%	2.5%	97.5%
	Medium Standard Deviation			
A $y \geq 90$	78.3%	21.7%	0.0%	0.0%
P $60 \leq y < 90$	21.7%	78.3%	78.3%	21.7%
F $y < 60$	0.0%	0.0%	21.7%	78.3%
	High Standard Deviation			
A $y \geq 90$	65.2%	34.8%	0.0%	0.0%
P $60 \leq y < 90$	34.8%	65.2%	65.2%	34.8%
F $y < 60$	0.0%	0.0%	34.8%	65.2%

Table 1 gives the probabilities of earning an A, Pass, or F when the study effort is such that the earned number of points are 91, 89, 61, and 59, that is, quite near the grade cutoff points.

It should first be said that these examples are meant to be suggestive only and other choices could certainly be made. The basic notion is that things happen. A student might for example write the due date for an important project on their calendar for a week after the actual due date. They devoted the time and did the project, but their grade is reduced, for example, by 5% of the semester grade because of the missed project deadline. On the other hand, a student might experience a positive random shock of 5% of the course grade if, for example, they are assigned to a group project which is well done and gets a good grade, but they did not even know they were assigned to that group and they did not contribute at all to the group project. (Other group members no doubt would experience this as a negative shock!) A student may decide to miss a class or two with the belief that further classes will not be missed, thus not affecting their final grade. However, the student may get sick or have another type of personal emergency and must miss additional classes thus decreasing their final grade. Some instructors offer extra credit that some students do not believe they need at that particular time, as they are confident in their study efforts. Some of these students will likely over estimate their study effort and earn a grade lower than expected.

To simplify the discussion of results we will call the “earned grade” the number of points a given amount of study effort would have earned in the absence of uncertainty. The examples in Table 1 are avowedly ones near the grade cutoffs. For students with earned grades near a cutoff there is a substantial probability of getting a course grade other than the one suggested by the earned grade. For students near but slightly above the cutoff, there is a 97.5%—65.2% chance of getting the earned grade, and a 2.5%—34.8% chance of getting the lower grade. The situation is just the opposite for person near but just below a cutoff grade. This student will have a 2.5%—34.8% chance of getting the higher grade and a 97.5%—65.2% of getting the earned grade.

Table 2 shows a similar analysis for a letter grade system with the grades A, B, C, D, and F with grade cutoffs of 90, 80, 70, and 60.

TABLE 2
LETTER GRADE PROBABILITIES

Earned Grade	89	87	85	83	81
	Low Standard Deviation				
A $y \geq 90$	2.5%	0.0%	0.0%	0.0%	0.0%
B $80 \leq y < 90$	97.5%	100.0%	100.0%	100.0%	97.5%
C $70 \leq y < 80$	0.0%	0.0%	0.0%	0.0%	2.5%
D $60 \leq y < 70$	0.0%	0.0%	0.0%	0.0%	0.0%
F $y < 60$	0.0%	0.0%	0.0%	0.0%	0.0%
	Medium Standard Deviation				
A $y \geq 90$	21.7%	0.9%	0.0%	0.0%	0.0%
B $80 \leq y < 90$	78.3%	99.1%	100.0%	99.1%	78.3%
C $70 \leq y < 80$	0.0%	0.0%	0.0%	0.9%	21.7%
D $60 \leq y < 70$	0.0%	0.0%	0.0%	0.0%	0.0%
F $y < 60$	0.0%	0.0%	0.0%	0.0%	0.0%
	High Standard Deviation				
A $y \geq 90$	34.8%	12.0%	2.5%	0.3%	0.0%
B $80 \leq y < 90$	65.2%	87.7%	95.0%	87.7%	65.2%
C $70 \leq y < 80$	0.0%	0.3%	2.5%	12.0%	34.8%
D $60 \leq y < 70$	0.0%	0.0%	0.0%	0.0%	0.0%
F $y < 60$	0.0%	0.0%	0.0%	0.0%	0.0%

For a given earned grade and standard deviation the result for the earned grade of 89 is the same as in Table 1. Similarly, since these are just probabilities the results for 81 are analogous to the results above for 61. Since the Pass grades are now divided into B, C, and D, we can see the probabilities for the range of grades for one of the middle grades, in the case shown for a B. For the lowest of the standard deviations, to one significant digit students with earned grades 87, 85, and 83 are going to get a B. Students at 89 have a 2.5% chance of an A while those at 81 have the same chance of a C.

Notice now that for the largest standard deviation students with an earned grade of 85 have a 2.5% chance of an A or C. Students at the top of this range with an earned 89 have a 34.8% chance of getting an A and still no chance of getting a C (to one significant digit) even with this relatively large standard deviation. With this largest standard deviation students with an earned grade of 87 have a 12% chance of getting an A and a barely perceptible chance (.3%) of getting a C. At the lower end of this grade range, 83 and 81, the probabilities are reversed.

REFLECTIONS FROM THE TEACHER'S DESK

Up to this point we have been looking at student effort from the point of view of the student, with the basic premise being that there is a gap between the grade earned by study effort and the actual number of points earned—that from the point of view of the student is random—outside of the student's control. Instructors are sensitive to this gap as well, but inevitably have different experiences and perceptions. At the most basic level, the relationship between effort and earned grade requires that a student know their ability, since ability level is a parameter in the study effort function. It may be that a student has not taken a class like this one before and has an imperfect grasp of their ability level relevant to the course at hand. If a student misestimates their own ability level by just a bit, a study effort that they thought would lead to a 91 actually generates an earned grade of 89. If such misapprehensions were symmetrical and unbiased then the opposite case would be as prevalent, they thought their earned grade would be 89 when it was in fact 91. No doubt both things happen, but the cognitive bias literature has identified two biases that might be relevant, the over-optimism bias and the over-precision bias. (For a succinct intuitive discussion of these see Pindyck and Rubinfeld, 2018, p. 700-701.) The former is the unrealistic belief that things will work out well and the latter is the unrealistic belief that one can predict outcomes.

And especially in this grading system there are significant differences between an 89 and a 91. (The next section will consider plus/minus grading.) A student who expects an 89 and gets a 91, has a "gain" of a couple of points that shifts the grade from B to A. Conversely the student who expects a 91 and gets an 89 has a loss of a couple of points that shifts the grade from A to B. So there is quite a lot objectively at stake in this size gain/loss. But additionally, it may be the case that students experience "loss aversion". This is the tendency of individuals to prefer avoiding losses as compared to acquiring gains. (Again see Pindyck and Rubinfeld, 2018, p. 693.) In fact, there is considerable evidence that many people mind a loss about twice as much as they value an equal size gain (see Kahneman, 2011, 284). If this rule of thumb holds for student grades, a student would be twice as upset for an 89 compared to the expected 91 than they would be pleased at a 91 when they expected 89.

When students later think back to reflect on the particular course, factors that are random from the students' point of view might well influence their impressions on what they gained from the course. To simplify the analysis our model assumes a single grading event, presumably the semester grade. But in practice (as reflected by many of the examples from our own teaching) the gap can occur before the end of the course, at a time when the student still has a chance to respond by changing their subsequent study effort. As an example, some institutions report midterm grades. Students may still have an opportunity to make a shift in their study effort in an attempt to change the ultimate final score/grade in the course. In fact, the purpose of the Oettinger (2002) paper from which our examples are adapted is to study empirically the question of whether students near but above a grade cutoff going into the final exam behave differently from those near but below the cutoff. His primary finding is that "students who are on a borderline between grades going into the final exam score higher on the final exam than do other students, after controlling for pre-final exam performance" (Oettinger, 2002, p. 517). (Note the contrast in findings with Grant and Green.) Interestingly he says that such students are attempting to "game" the grading system, and suggests a plus/minus system might reduce such gaming, again, an issue we will return to shortly. As instructors at an institution without pluses or minuses, we have experienced students discussing strategies to "game" the system to ultimately achieve a desired score without strenuous study effort by the students. Students will spend a considerable amount of time and effort calculating the minimum number of points need on a final exam in order to achieve their desired final grade.

Regardless of the grading system and where the student is in the course, a gap between earned grade and actual grade that is random from the student's point of view might not be random from the instructor's point of view. Tables 1 and 2 present some basic probabilities in a systematic way, assuming a mean and standard deviation in random points. The assumption is that students cannot affect that probability distribution in any way—they have to take it as given. But just because students cannot impact the probability distribution does not mean that instructors are powerless. To say the same thing in different words, reflecting on cases in their own classes where students seems to have had a gap between effort and course grade might cause an instructor to wonder whether there was anything the instructor might have done differently. In particular, instructors might wonder whether their actions could change either the mean or the standard deviation of the probability distribution.

Changing the mean might in fact be possible. Although it's more complicated than this in real life, to oversimplify an instructor might have an unexpected extra credit question on the final exam with an expected grade of 1 (e.g. a question might be worth 2 points and on average half of the students get it right). They might find some other strategy to simply add 1 point to everyone's grade. (If they teach the same course and/or students over and over it needs to be something that the students don't come to expect—a certain amount of creativity might be required.) This would have the effect of making the random distribution have a mean of +1 with an unchanged (or essentially unchanged) standard deviation.

Among the benefits of such a strategy is its simplicity and relative ease. A person might however reflect on the effect on the overall class GPA, as it certainly would raise it, although by a slight amount.

It might also be possible to leave the mean unchanged and reduce the standard deviation. Many of the examples in this paper, drawn from our experiences and the experiences of our colleagues, are truly random—nothing could be done to change them. But grading often involves an irreducible subjective element, and instructor behavior might affect the distribution of grades. For example, consider the challenges of assigning partial credit on an essay exam. There is always the risk that the first few papers are graded more strictly or more easily than the last few. The purely random factor of where a student's paper is in the stack might affect its grade. To guard against this, many instructors routinely go back and re-grade those at the top of the stack, but the extent to which this is done might well affect the overall standard deviation of the random effects. Although dispersion can never be completely eliminated, sufficient re-grading might considerably reduce the standard deviation.

It should quickly be said that efforts to reduce the standard deviation always have an opportunity cost. Time spent obsessing over exam grading might be better spent preparing for the next class section or making notes for the next time the course is taught. Instructors also are aware of their own tendencies to favor some students. Anonymous grading would avoid this tendency. Students may experience tragedies that may tug on the emotions of the instructor. However, the instructor should understand that fairness and justice do not allow for grade deviations. Disruptions may occur during lectures, presentations, or exams that may affect the overall effort in the class. Another reality is that many pre-tenure instructors must be thinking about the affect grades and student effort will have on faculty evaluations, and ultimately their tenure decision.

REFLECTIONS ON THE CHOICE OF A GRADING SYSTEM

Using a letter grade system rather than a strict numerical system creates the problem of choosing grade cutoffs. The discrete grade difference between a grade of 89 and 90, if 90 is the cutoff, causes much anguish and consternation for both students and teachers. Individual instructors often choose the grade cutoffs for a particular course. However, the grading system itself, whether A/Pass/Fail, just letter grades, or a plus/minus system, is typically chosen by a committee of instructors rather than a single instructor. The obvious difference between these is the number of cutoffs that must be chosen. A quick look at the literature suggests there are a number of possibilities. One such assumes the grades are A, A-, B+, B, B-, C+, C, C-, D, and F with the cutoffs 94, 90, 87, 84, 80, 77, 74, 70 and 65. Table 3 gives the results for earned grades in the 80s assuming the medium standard deviation from above.

TABLE 3
PLUS/MINUS GRADES PROBABILITIES

E.G.	89	88	87	86	85	84	83	82	81	80
A-	21.6%	5.8%	0.9%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
B+	72.5%	72.5%	49.1%	21.6%	5.8%	0.9%	0.1%	0.0%	0.0%	0.0%
B	5.8%	21.6%	49.1%	72.5%	72.5%	49.1%	21.6%	5.8%	0.9%	0.1%
B-	0.0%	0.1%	0.9%	5.8%	21.6%	49.9%	77.4%	88.3%	77.4%	49.9%
C+	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.9%	5.8%	21.6%	49.1%
C	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.9%

In this system an actual grade of 89, 88, and 87 gets a B+, 86, 85, and 84 a B, 83, 82, 81, and 80 gets a B- minus. (The grades of 90, 91, 92, and 93 get an A- and the grades of 79, 78, and77 get a C+, etc. The table only includes rows for grades with a positive probability when measured to one significant digit.) Notice that more grades means that the 10 numbers in the 80s have to be divided into three categories, and if grades are rounded to whole numbers, then two grades will have three entries, and one four. This slight asymmetry changes the probabilities slightly for earned grades exactly at cutoff. (In the basic grading system, a student getting an 80 would have a 50% chance of getting a B and a 50% change of getting a C; the random component is just as likely to put the student slightly above 80 at it is below 80 and there is 0% of exactly 80 because of the continuous probability distribution.)

Looking at Table 3, notice that a student with an earned grade of 80 has a 49.9% chance of a B- and a very slight (.1%) chance of a B. That same student has a 49.1% chance of getting a C+ and a slight (.9%) chance of getting a C. A student exactly at one of the other cutoffs (87, 83) has the same probabilities. A student with a grade one below a cutoff for the B+ and B grades (89, 86) has a 72.5% chance of getting that grade, a 21.6% chance of the higher grade and a 5.8% chance of the lower. For a grade two below a cutoff of these grades (88, 85) the probabilities are reversed, that is, there is still a 72.5% chance of getting the given grade but a 5.8% chance of the higher grade and a 21.6% chance of the lower one. For the B- grade the probabilities are slightly different. For 83 and 81 the probability of getting a B- is 77.4%. For a 83 there is a very slight (.1%) chance of a B+, a 21.6% chance of a B and a slight (.9%) chance of a C+. For an 81 the probabilities are reversed. This grading system assigns the one extra grade to the B- group; it is 82, which has an 88.3% probability of a B- and a 5.8% chance of a B and an equal chance of a C+.

Looking at grades in such a granular way seems odd in some sense. But our experience as instructors suggests that it is not at all uncommon for students and instructors to look at grades this carefully. One kind of thinking about grading sees the larger number of grades in a plus/minus system as a positive because it reduces what is at stake. The difference between a 90 and an 89 is the difference between an A- and a B+ rather than an A and a B. This might be related to Oettinger's reflection that a plus/minus system might reduce the extent to which students try to game the grading system because there is less to be gained, other things being equal. It would seem that both student and instructor stress would also be less in this situation. The other side of that argument is that it creates many more cutoff points, so the number of students sitting at a grade cutoff increases substantially. Perhaps student effort will be higher.

A back of the envelope definition of gaming the grading system might be that a student devotes time to lobbying for a grade that they might have devoted to study and increasing the earned grade. A plus/minus system reduces the incentive to lobby if you land on (or near) a cutoff, but increases the number of cutoffs. It is nothing like evident what the impact on student time allocation decisions would be. Another example of gaming the system with the A-B-C-D-F grading system is that students often try to compute their minimum study effort for final exam. Students realize that some effort are unrealistic and focus their efforts in other courses where increased study effort will gain them the most successes.

FINAL THOUGHTS

An individual instructor controls some aspects of syllabus design but not others. For example, the grading system (A/Pass/Fail, A-F, plus/minus) is a choice, but not typically one made by individual instructors. The objective of the group choosing the grading system would presumably include generally student welfare, but it is not self-evident that increasing learning as measured by a grade in a given course is the primary goal. Grant and Green (2013) note the history of changing the grading convention from numerical marks to letter grades at Harvard College in the 1880s was to "diminish the competition for marks and the importance attached by students to College rank in comparison with the remoter objects of faithful work." (Please see Grant and Green (2013, p. 1565) for the original source of this quotation.) Grant and Green argue that the change to a letter grade system "was explicitly intended to *diminish* motivation" (2013, 1565, italics in the original), an inference that we would not necessarily draw.

The use of a particular grading system may implicitly encourage time spent on non-graded activities essential to what it means to go to college. For example, students that waste time calculating how to improve from a 96 to a 98 score may be better served by the utilization of their time and efforts in the overall college experience. As an example, one instructor offers two extra credit points for students to attend various on-campus speakers related to the course content, such as a U.S. Supreme Court Justice.

Students often attend these experiences partially in order to receive these points. However, if the course has 400 points, these two-point extra credit experiences do not really change their grade unless the student is so close to the next grade cut off. The college experience is really about the overall effort, not just the effort in a particular course. Students should be exposed to various points of view, particularly a point of view not exactly like their own.

REFERENCES

Allgood, S., Walstad, W.B.,& Siegfried, J.J. (2015). Research on Teaching Economics to Undergraduates. *Journal of Economic Literature,* 53(2), 285-325.

Babcock, P. (2010). Real Costs of Nominal Grade Inflation? New Evidence from Student Course Evaluations. *Economic Inquiry,* 48(4), 983-996.

Babcock, P.,& Marks, M. (2011). The Falling Time Cost of College: Evidence from Half a Century of Time use Data. *Review of Economics and Statistics,* 93(2), 468-478.

Bonesronning, H.,& Opstad, L. (2015). Can Student Effort be Manipulated? Does it Matter? *Applied Economics,* 47(13-15), 1511-1524.

Bonesronning, H.,&Opstad, L. (2012). How Much is Students' College Performance Affected by Quantity of Study? *International Review of Economics Education,* 11(2), 46-63.

Brustin, S.L.,& Chavkin, D.F. (1997). Testing the Grades: Evaluating Grading Models in Clinical Legal Education*, Clinical Law Review*, 3, 299-332.

Butcher, K.F., McEwan, P.J.,& Weerapana, A. (2014). The Effects of an Anti-Grade-Inflation Policy at Wellesley College. *Journal of Economic Perspectives,* 28(3), 189-204.

Charlton Jr., W. T.,& Stevens, J.L. (2000). Searching for Excellence among College Undergraduates: Applications of Financial Analysis Principles to Student Transcripts. *Financial Practice and Education*, 10(2), 175-181.

Grant, D.,& Green, W.B. (2013). Grades as Incentives, *Empirical Economics,* 44(3), 1563-1592.

Forsyth, B., Kimble, C., Birch, J., Deel, G., &Brauer, T. (2016). Maximizing the Adaptive Learning Technology Experience. *Journal of Higher Education Theory and Practice,* 16(4), 80-88.

Hamer, L.O.,& O'Keefe, R.D. (2013). Achieving Change in Students' Attitudes Toward Group Projects. *Journal of Higher Education Theory and Practice,* 13(2), 25-33.

Hernandez-Julian, R.,& Looney, A. (2016). Measuring Inflation in Grades: An Application of Price Indexing to Undergraduate Grades. *Economics of Education Review,* 55, 220-232.

Ho, D. E.,&Kelman, M.G. (2014). Does Class Size Affect the Gender Gap? A Natural Experiment in Law. *Journal of Legal Studies,* 43(2), 291-321.

Kahneman, D. (2011). *Thinking Fast and Slow*. NY: Farrar, Straus, and Giroux.

Oettinger, G.S. (2002). The Effect on Nonlinear Incentives on Performance: Evidence from Econ 101. *Review of Economics and Statistics* 84(3), 509-517.

Perlman, B., McCann, L.,& Prust, A. (2007). Students' grades and ratings of perceived effectiveness of behaviors influencing Academic Performance. *Teaching of Psychology,* 34, 236-240.

Peterson, C.M.,& Peterson, T.O. (2016). Affecting Student Perception and Performance by Grading With 10,000 Points. *Journal of Management Education,* 40(4), 415 –431.

Pindyck, R.S.,& Rubinfeld, D.L. (2018). *Microeconomics* 9th, ed. Pearson.

Rose, L.M. (2011). Norm-referenced Grading in the Age of Carnegie: Why Criteria-Referenced Grading is More Consistent with Current Tends in Legal Education and How Legal Writing can Lead the Way. *Legal Writing: J. Legal Writing Inst.*, 17, 123-159.

Stinebrickner, R.,&Stinebrickner, T.R. (2008). The Causal Effect of Studying on Academic Performance. *B.E. Journal of Economic Analysis and Policy: Frontiers of Economic Analysis and Policy,* 8(1), 1-55.

APPENDIX

Consider the model of Oettinger (2002) where student "performance," p, (the example being the proportion of total possible points earned) is a production function $p = f(e,a)$ of the choice variable "study effort," e, the student's endowment of "ability," a. The number of points earned is $y = p + u = f(e,a) + u$. That is, expected student performance is subject to a random shock, u, with a continuous unimodal density function, with the mode at zero. The resulting cumulative distribution function is denoted $G(u)$. These assumptions allow the determinate performance p to be an unbiased predictor of course grade. There are a discrete number of course grades, Y_i, each of which generates a discrete utility, U_i. For example, a student receives grade i when $Y_{i-1} > y \geq Y_i$ or $Y_{i-1} - f(e,a) > u \geq Y_i - f(e,a)$ and receives a utility of U_i. (Y_{i-1} is a higher grade than Y_i and $G(Y_0 - f(e,a)) = 1$ and $G(Y_n - f(e,a)) = 0$.) Students are purely grade motivated in the since that utility depends on the grade earned rather than on the knowledge gained. The student's optimization decision is to maximize expected utility given that there is a cost of providing study effort $C(e)$.

In the very simple example of three grades A/P/F with cutoffs at 90% for an A and 60% for a pass so that $Y_0 = \infty$ $Y_1 = 90$, $Y_2 = 60$, and $Y_3 = 0$ and grade 1 is A, grade 2 is P, and grade three is F. Please see Table A.

TABLE A
A, PASS, FAIL

	Earned Grade $f(e,a) = 87$			
	$Y_0 = \infty$	$Y_1 = 90$	$Y_2 = 60$	$Y_3 = 0$
	$\Pr(y \geq Y_1)$ $= 1 - G(Y_1 - f(e,a))$		$\Pr(y \geq 90)$ $= 1 - G(90 - 87)$ $= 1 - G(3)$	
	$\Pr(Y_1 \geq y \geq Y_2)$ $= G(Y_1 - f(e,a))$ $- G(Y_2 - f(e,a))$		$\Pr(90 \geq y \geq 60)$ $= G(90 - 87) - G(60 - 87)$ $= G(3) - G(-27)$	
	$\Pr(y \leq Y_3)$ $= G(Y_2 - f(e,a))$ $- G(Y_3 - f(e,a))$		$\Pr(y \leq 60)$ $= G(60 - 87) - G(0 - 87)$ $= G(-27) - G(-87)$	

Consider the probability of getting an A.

$$\Pr(y \geq Y_1(=90)) = \Pr(f(e,a) + u \geq Y_1) = \Pr(u \geq Y_1 - f(e,a)) = 1 - G(Y_1 - f(e,a)).$$

That is, the probability that the score is 90 or better is equal to the probability that the cutoff minus the earned grade is less than the random variable. E.g. suppose $f(e,a) = 87$ so that $Y_1 - f = 90 - 87 = 3$. Then the probability that the score is 90 or better is equal to the probability that the random number of points earned is 3 or more, which is one minus the cumulative probability of three or fewer points. The contribution to the expected grade is $U_1 * (1 - G(Y_1 - f(e,a)))$. To get a passing grade the random number of points gained must less than 3, and you can't randomly lose more than 27 points (i.e., $Y_2 - f = 60 - 87 = -27$); recall the "points gained" can be negative. The probability of a random number of points being less than 3 and greater than -27 is $\Pr(Y_1 \geq y \geq Y_2) = \Pr(60 \leq y \leq 90) = G(Y_1 - f(e,a)) - G(Y_2 - f(e,a))$. (The distribution function is continuous so the probability of exactly -27 points is zero.) The contribution to the expected grade is $U_2 * G(Y_1 - f(e,a)) - G(Y_2 - f(e,a))$. Finally, to get a failing grade the number of random points gained must be less than -27, that is, the number of random points lost must be greater than 27.